

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



**The Strange Stories Film Task
A New Measure of Social Cognition.**

Murray, Kim

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Volume I
Main Research Project
Service Evaluation

Mr Kim Murray

**Thesis submitted in the partial fulfilment of the degree of Doctor of
Clinical Psychology**

**Department of Psychology, Institute of Psychiatry, King's College
London**

2014

Acknowledgements

It has been a hard and rewarding journey and I would like to thank my friends and colleagues who have supported me throughout the work presented in this volume. Firstly, for the main research project I would like to thank Prof. Francesca Happé who has always shown kindness and support and her passion for research is infectious. Dr Kate Johnston, whom I have worked with for many years now, has provided invaluable insights from the clinical and academic arenas throughout this work. As my personal support tutor I would also like to thank her for her guidance regarding career more generally.

I have Ms Helen Cunane to thank for bringing the project to life and editing the films with such care and enthusiasm. The two actors who volunteered their time and were willing to bare my direction, I would also like to thank for their patience and generosity. While the project has developed and evolved since, I would like to thank Ms Charlotte Wormald and Dr Ailsa Russell for their invaluable contributions to the early phases of the project and its inception. Similarly, Sean Taylor's invaluable contributions during the early phases of the project were extremely helpful. I am indebted to the BGC and Outpatient ASD team at the Maudsley Hospital who facilitated data collection and my thanks goes to Dr Neil Hammond and Ms Jenna Vyass in particular. I would like to thank Dr Tim Cadman for his proof reading of the document and his knowledge of ROC curves.

I am grateful to Dr Georgina Krebs' ambition and support with the service evaluation project and her willing to go 'above and beyond' to support my learning. I would like to thank Dr Amita Jassi, Ms Faye Barrows and Prof. David Mataix-Cols for their patience and constructive feedback when writing up this work.

My fellow trainees have been an unforgettable melody of support, humour and comradery throughout this process and I would like to thank them whole heartedly for their indirect contribution to this project. I would also like to thank Keerthana for her patience, relaxing presence and kindness.

Finally, I would like to thank all the participants who volunteered their time to enable this work to come to fruition.

Table of Contents

| | |
|---|------------|
| Main Research Project: The Strange Stories Film Task: A New Measure of Social Cognition..... | 4 |
| Service Evaluation: . Referral Rates, Clinical Presentations and Outcomes of Cognitive Behaviour Therapy for Obsessive-Compulsive Disorder in Young People with and Without Autism Spectrum Disorders..... | 212 |

The Strange Stories Film Task: A New Measure of Social Cognition.

Mr. Kim Murray

Supervised by Prof. Francesca Happé and Dr Kate Johnston

Main Research Project

Department of Psychology, Institute of Psychiatry

King's College London

Abstract

Introduction: Autism Spectrum Disorders (ASD) are characterised by difficulties in social interaction. High functioning (HF) adults with an ASD diagnosis often report subtle social cognitive difficulties. The main aim of the study was to develop and validate a novel measure of social cognition (The Strange Stories Film Task (SSFt)) and in doing so overcome a number of limitations to available measures in the field.

Method: The measure consisted of acted scenarios designed to capture the subtle mentalizing difficulties observed in adults with high functioning ASD. 20 participants were recruited to pilot the new measure. A final test set was produced and shown to a group of 20 well diagnosed HFASD adults and matched controls. Participants also completed well established measures of social cognition and questionnaire measures of empathy, alexithymia and ASD traits.

Results: The SSFt was more effective at differentiating the HFASD group from the control group showing greater levels of sensitivity. Group differences could not be attributed to general cognitive factors. The SSFt was associated with the traditional measures of social cognition. Performance on the SSFt was associated with measures of empathy and ASD symptomatology. No associations with alexithymia were observed.

Conclusion: The SSFt is a potentially useful tool to identify mentalizing difficulties in HFASD samples. In addition, the measure was sensitive to individual differences in mentalizing abilities in non-autistic adults. The SSFt showed adequate convergent validity. The elements of the measure targeting social interaction abilities rather than understanding proved the most sensitive. These findings are discussed with regard to clinical implications and future research.

Table of Contents

| | |
|---|-----------|
| List of Tables..... | 10 |
| List of Figures | 10 |
| Abbreviations..... | 11 |
| 1 Introduction..... | 13 |
| 1.1 Autism Spectrum Disorders..... | 13 |
| 1.1.1 Autism Spectrum Disorder Diagnosis and Terminology..... | 13 |
| 1.1.2 Prevalence and Aetiology..... | 14 |
| 1.1.3 Cognitive Accounts of ASD..... | 15 |
| 1.2 Theory of Mind..... | 16 |
| 1.3 Assessing/Measuring Theory of Mind..... | 17 |
| 1.3.1 False Belief Tasks | 17 |
| 1.3.2 Overcoming Initial Problems with False Belief Tasks..... | 19 |
| 1.3.3 Mentalizing and Emotion Recognition: More Ecologically Valid? | 22 |
| 1.4 Ecologically Valid Tasks of Social Cognition..... | 23 |
| 1.4.1 Empathic Accuracy | 24 |
| 1.4.2 Social Cognition Measures using Dynamic Stimuli..... | 25 |
| 1.5 Cognitive vs. Affective Empathy..... | 29 |
| 1.6 The Alexithymia Hypothesis..... | 32 |
| 1.7 Successful Mentalizing Equals Appropriate Social Interaction?..... | 34 |
| 1.8 Aims of the Current Study | 35 |
| 1.9 Hypotheses | 36 |
| 1.9.1 Group Differences..... | 36 |
| 1.9.2 Convergent Validity | 36 |
| 1.9.3 Empathy | 36 |
| 1.9.4 Alexithymia | 37 |
| 2 Method..... | 38 |
| 2.1 Devising the Strange Stories Film Task (SSFt)..... | 38 |
| 2.1.1 Scripts..... | 38 |
| 2.1.2 Actors..... | 39 |
| 2.1.3 Filming and Editing..... | 39 |
| 2.1.4 Questions | 40 |
| 2.1.5 Order..... | 42 |

| | | |
|-------|---|----|
| 2.2 | Piloting | 43 |
| 2.2.1 | Ethics | 43 |
| 2.2.2 | Participants | 43 |
| 2.2.3 | Procedure | 44 |
| 2.3 | Measures..... | 44 |
| 2.3.1 | The Autism Quotient (AQ) | 44 |
| 2.3.2 | The Strange Stories Film Task Pilot..... | 45 |
| 2.4 | Selection of Final Test Set..... | 45 |
| 2.4.1 | Scene Selection | 45 |
| 2.4.2 | Order effects | 46 |
| 2.4.3 | Imposing a Narrative | 46 |
| 2.4.4 | Prompting..... | 47 |
| 2.5 | Scoring..... | 47 |
| 2.5.1 | Experimental Clips | 47 |
| 2.5.2 | Control clips | 49 |
| 2.6 | Psychometrics | 50 |
| 2.6.1 | Item Analysis | 50 |
| 2.6.2 | Inter-rater Reliability..... | 51 |
| 2.7 | Phase 2: | 51 |
| 2.7.1 | Ethics | 51 |
| 2.7.2 | Design..... | 52 |
| 2.7.3 | Participants | 52 |
| 2.7.4 | Inclusion Criteria | 52 |
| 2.8 | Measures..... | 54 |
| 2.8.1 | Wechsler Intelligence Tests | 54 |
| 2.8.2 | The Autism Quotient (AQ) | 55 |
| 2.8.3 | The Twenty item Toronto Alexithymia Scale (TAS-20)..... | 55 |
| 2.8.4 | The Interpersonal Reactivity Index (IRI) | 55 |
| 2.8.5 | Reading the Mind in the Eyes task (Eyes) | 56 |
| 2.8.6 | The Awareness of Social Inference Test (TASIT) | 57 |
| 2.8.7 | The Frith-Happé Triangles (F-HT) | 57 |
| 2.8.8 | The Strange Stories (SS) | 58 |
| 2.8.9 | The Strange Stories Film Task(SSFt)..... | 59 |
| 2.9 | Procedure..... | 60 |
| 2.10 | Statistical Analysis..... | 60 |

| | | |
|----------|---|------------|
| 3 | Results | 62 |
| 3.1 | Traditional Social Cognition Measures | 63 |
| 3.2 | Questionnaire Measures | 65 |
| 3.3 | The Strange Stories Film Task..... | 67 |
| 3.4 | Group Classification Using the SSFt and Traditional Social Cognition tasks...73 | |
| 3.5 | Convergent Validity of the Strange Stories Film Task. | 75 |
| 3.6 | The Strange Stories Film Task's Relationship with ASD Traits/Symptoms | 77 |
| 3.7 | Summary of Main Findings | 78 |
| 4 | Discussion..... | 79 |
| 4.1.1 | Group Differences..... | 79 |
| 4.1.2 | Convergent Validity | 81 |
| 4.1.3 | Empathy | 84 |
| 4.1.4 | Alexithymia | 85 |
| 4.2 | The Strange Stories Film Task | 86 |
| 4.2.1 | Understanding the Intention question | 86 |
| 4.2.2 | Verbal Abilities and the Strange Stories Film Task | 87 |
| 4.2.3 | Understanding the Interaction Question..... | 88 |
| 4.3 | Limitations of the study..... | 92 |
| 4.3.1 | Generalisability | 92 |
| 4.3.2 | The Strange Stories Film Task..... | 93 |
| 4.4 | Clinical Implications..... | 95 |
| 4.5 | Future Research..... | 96 |
| 4.6 | Concluding Remarks..... | 99 |
| 5 | References..... | 100 |
| 6 | Appendices..... | 121 |
| 6.1 | SSFT Experimental Scripts | 121 |
| 6.2 | SSFT Control Scripts | 138 |
| 6.3 | SSFT Project Brief for Actors..... | 143 |
| 6.4 | Pilot Information Sheet and Consent Form..... | 145 |
| 6.5 | The Autism Quotient..... | 148 |
| 6.6 | SSFT Pilot Instructions | 152 |
| 6.7 | Scene Selection..... | 162 |
| 6.8 | SSFT Scoring..... | 165 |
| 6.9 | Information Sheets and Consent Forms for SSFt Phase 2 | 182 |
| 6.10 | The TAS-20 | 192 |

| | | |
|------|--|-----|
| 6.11 | The IRI | 194 |
| 6.12 | Strange Stories and Scoring..... | 197 |
| 6.13 | SSFt instructions, Scoring Sheet and Order A (Order B is reversed) | 203 |
| 6.14 | Associations Between Traditional ToM Measures..... | 208 |
| 6.15 | Traditional ToM Measures and IRI, TAS-20,ADI-R and AQ Associations. . | 209 |
| 6.16 | SSFt Order Effects..... | 211 |

List of Tables

| | |
|--|----|
| Table 1: Characteristics of current dynamic social cognition tasks. | 26 |
| Table 2: Internal consistency for both the Experimental and Control film clips. | 50 |
| Table 3: Intra-class coefficients for each dimension of the SSFt for both the ASD and control group..... | 51 |
| Table 4: Demographics for both groups: Mean (SD) | 53 |
| Table 5: Results by group for tests of social cognition: Mean (SD) | 63 |
| Table: 6: Self- reported empathy and alexithymia: Mean (SD) | 65 |
| Table 7: Performance on SSFt by group: Mean (S.D.)..... | 67 |
| Table 8: Cut-off scores for each social cognition measure. | 74 |
| Table 9: Associations by group Between Performance on the SSFt and Traditional Social Cognition Measures: Spearman's partial correlation coefficient controlling for VIQ (r^2). | 75 |
| Table 10: Association of the SSFt with alexithymia and empathy: Pearson's partial correlation coefficients after controlling for VIQ (r^2)..... | 76 |
| Table 11: Association of SSFt and ASD symptoms: correlation coefficient after controlling for VIQ (r^2)..... | 77 |
| Table 12: Summary of findings across the social cognition measures..... | 78 |

List of Figures

| | |
|---|----|
| Figure 1: Sum of scores on the SSFt by theme in Control group. Presented in descending order of Intention score for the ASD group | 69 |
| Figure 2: Sum of scores on the SSFt by theme for the ASD group presented in descending order of Intention score. | 69 |
| Figure 3: Difference scores between control and ASD group on the SSFt by theme in descending order of Intention scores for ASD group..... | 71 |
| Figure 4: Relationship of performance on the Intention and Interaction questions of the SSFt for both groups | 72 |
| Figure 5: ROC curves for each social cognition measure and chance..... | 73 |

Abbreviations

AaB: Above and Beyond Task

ADI-R : Autism Diagnostic Manual-Revised

ADOS: Autism Diagnostic Observation Schedules

AMT: Awkward Moments Test

AQ: Autism Quotient

AS: Asperger's syndrome

ASD: Autism Spectrum Disorders

ATOMIC: The Animated Theory of Mind Inventory for Children

CC: Central Coherence

CI: Confidence Intervals

D: Cohen's D

Df: Degrees of freedom

DSM-5: Diagnostic and Statistical Manual of Mental Diseases 5th edition

DSM-IV: Diagnostic and Statistical Manual of Mental Diseases 4th edition

EC: Empathic Concern

EF: Executive Function

EQ: Empathy Quotient

ER: Emotion Recognition

Eyes: Reading the Mind in the Eyes Task

F: F- value.

F: Fantasising

FC: Forced choice

F-H T: Frith Happé triangles

HFASD: High Functioning Autism Spectrum Disorders

ICD-10 : International Classification of Diseases, 10th edition

IQ: Intelligence Quotient

IRI: Interpersonal Reactivity Index

MASC: A Movie for the Assessment of Social Cognition

MASC-FC: A Movie for the Assessment of Social Cognition- Forced Choice version.

MASC-SP: A Movie for the Assessment of Social Cognition- Spanish Version

MDFT: Moral Dilemmas Film Task

PD: Personal Distress

PDD: Pervasive Developmental Disorders

PST: Psychological State Talk

PT: Perspective Taking

RMFT: Reading the Mind in the Films Task

SP: Spanish

SS: Strange Stories

SSFt: Strange Stories Film Task

SSFt-p: Strange Stories Film Task Pilot

t: t-statistic

TAS-20: Twenty item Toronto Alexithymia Scale

TASIT: The Awareness of Social Inference Test

ToM: Theory of Mind

VIQ: Verbal IQ

WAIS-III: Wechsler Adult Intelligence Scale –III

WAIS-IV: Wechsler Adult Intelligence Scale –IV

WASI: Wechsler Abbreviated Scale of Intelligence

1 Introduction

1.1 Autism Spectrum Disorders

1.1.1 Autism Spectrum Disorder Diagnosis and Terminology

The Diagnostic and Statistical Manual of Mental Diseases 4th edition (DSM-IV) (American Psychiatric Association, 2000) and the International Classification of Diseases, 10th edition (ICD-10) (World Health Organization, 1992), use the terminology ‘triad of impairments’; difficulties in social interaction, communication and stereotyped behaviours (Wing, 1981; Wing & Gould, 1979), to describe Autism Spectrum Disorders (ASD). For a diagnosis to be made, symptoms must be present by the age of three. Asperger’s syndrome (AS) and autism are considered separate categories under the umbrella term of pervasive developmental disorders (PDD). Within these classification systems it is the presence (or lack thereof) of delayed cognitive and linguistic development that distinguishes childhood autism from AS.

In 2013, the DSM-IV was revised and updated. Within DSM-5 (American Psychiatric Association, 2013) the term pervasive developmental disorder (PDD) has been replaced with an overarching category of ASD that collapses the diagnostic sub-categories into a single spectrum. The notion of a diagnostic continuum, being of greater explanatory power than a categorical system, has been favoured by some (Murphy, Beecham, Craig, & Ecker, 2011; Wing, Gould, & Gillberg, 2011) while criticised by others (Tsai & Ghaziuddin, 2014). Of particular weight in favour of using a dimensional, rather than categorical, system is concern about the reliability with which information regarding severity of symptoms across the triad, IQ and linguistic abilities is used across different sites for diagnosis (Lord, Petkova, Hus, & et al., 2012). In their multi-site study, Lord, et al. (2012) found that the distribution of autistic symptomatology as measured by the Autism Diagnostic Observation Schedules (ADOS; Lord et al., 1989) and the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Couteur, 1994) were equivalent across sites. However, the distribution of diagnostic categories assigned varied substantially across site, despite each showing high levels of expertise. These findings suggest that the differences in categorisation are more influenced by clinician factors rather than standardised (and to some extent more objective) measurements.

In contrast, Tsai and Ghaziuddin (2014) have argued in their review, that the observed differences between autism and AS over a range of studies questions their integration into a unitary spectrum. These changes have a number of important effects including questions regarding validity (McPartland, Reichow, & Volkmar, 2012), impact on prevalence (Matson, Hattier, & Williams, 2012), and sociological impact e.g. the effects on the label AS being removed on the identity of such individuals (Wing, et al., 2011).

For the remainder of this review the terms high functioning (HF) autism and AS will be used under the umbrella term of HFASD to reflect this change in thinking.

1.1.2 Prevalence and Aetiology

Rates of ASD have been shown in epidemiological research to be approximately 6-10 in 1000 (Simonoff et al., 2008) and the gender ratio of male to females is approximately 3:1 (Baird et al., 2006). Some theorists have considered ASD as being best described as an 'extreme male brain' (Baron-Cohen, 2002), however the notion of a 'female protective effect' has also been evidenced (Murphy, et al., 2011; Robinson, Lichtenstein, Anckarsäter, Happé, & Ronald, 2013).

Despite ASD being behaviourally defined, there is a consensus that the disorder is biologically-based. Twin studies yield concordance rates for monozygotic twins that are extremely high (between 70-90%) and for dizygotic twins the rate is substantially less, 0-25% (Trikalinos et al., 2005). A recent meta-analysis suggests that environmental factors may account for more variance than previously expected (Hallmayer, Cleveland, Torres, & et al., 2011). The biological argument is furthered by the presence of a 'broader autistic phenotype', whereby autistic traits that are also present in the general population (Happé, Ronald, & Plomin, 2006) are elevated in family members of those diagnosed (Sasson et al., 2013). Interestingly, this supports Kanner's (1943) initial observation regarding the characteristics of the cases he described, being reflected in the parents' presentations or personality (e.g. obsessionality). Advances in behavioural genetics have enabled the genetic underpinnings of the phenotypic heterogeneity observed in ASD to be better understood, for review see (Jeste & Geschwind, 2014). Similarly, advances in imaging technology have enabled structural differences in brain anatomy (Ecker et al., 2010) and

atypical functional connectivity associated with the condition to be identified (Philip et al., 2012, Travers et al., 2012)

1.1.3 Cognitive Accounts of ASD

There have been a number of cognitive accounts of ASD proposed over the last 30 years. The three that will be briefly discussed are the executive function (EF; Hill, 2004), 'central coherence' (CC; Happé & Frith, 2006) and theory of mind (ToM; Frith, Morton, & Leslie, 1991) accounts. Executive functions are a broad range of higher-order processes including updating, set-shifting, inhibition (Miyake et al., 2000) and working memory (Baddeley, 1992). Deficits in these processes have been shown in individuals with ASD (Hill & Bird, 2006) and some theorists argue that EF capabilities are necessary for later ToM capacities to develop (Pellicano, 2012). CC, is the tendency to process information as an integrated whole rather than as compartmentalised segments and the ability to integrate contextual information to form meaning (Happé & Frith, 2006). In ASD a 'weak CC' processing style affords a keen 'eye for detail' and has been linked to specialist abilities observed in ASD (Happé & Vital, 2009) although it may also impair social reasoning where integration of multiple cues guides interpretation of the social exchange (Vermeulen, 2014)

Initially research attempted to conceptualise the difficulties of people with ASD within a single account (Baron-Cohen, 2002; Russell, 1997), however recent developments from population level research have led some researchers to consider the triad of impairments best conceptualised as distinct and separable from one another (Happé, et al., 2006). Each cognitive theory therefore, may be unable to explain the entire constellation of ASD presentations (as it is not a unitary construct), but have differential explanatory power for the separate elements of the triad (for in-depth discussion on this issue see Brunsdon & Happé, 2014)

The ToM account will be the focus of the remainder of this review. Of all the cognitive theories it has received the most attention with regard to explaining the atypical interpersonal style seen as fundamental to ASD. How ToM is defined and more importantly how it is *measured* will be the focus of the remaining literature review with an

emphasis on recent measures that attempt to quantify subtle ToM difficulties within the HFASD population.

1.2 Theory of Mind.

ToM was traditionally defined as the ability to ‘impute[s] mental states to himself and to others’ (Premack & Woodruff, 1978, p515) and was termed a ‘theory’ due to the fact that mental states are both unobservable and one can infer or predict behaviour from them (Premack & Woodruff, 1978). ToM was first used in primate research and is still an area for current research (Call & Tomasello, 2008). Experimental work with humans has stimulated argument as to whether to regard ToM development as evidence for a domain-specific ‘modularity of mind’ hypothesis (Fodor, 1981; Scholl & Leslie, 1999) or alternatively to see ToM as a construct influenced by domain-general processes such as EF (Pellicano, 2010, 2012), emerging over the course of development as a specialised cognitive function (Karmiloff-Smith, 2009).

The definition of ToM, as above, has been broadened by theorists in the field of developmental psychology and researchers now use terms such as mentalizing, mind-reading or more broadly, social cognition (Dziobek et al., 2006). These newer definitions take into account the multi-faceted nature of the social world where integration of language, expression and, importantly, contextual information are required as human beings do not interact according to ‘if x then always y’ rules (Vermeulen, 2014). Such definitions also acknowledge the tension between ‘purely’ tapping a construct such as ToM and the additional cognitive demands of ever increasing ecologically valid tasks (Heavey, Phillips, Baron-Cohen, & Rutter, 2000). When a broader definition is required for the remainder of the review, ‘mentalizing’ or social cognition will be adopted to reflect this trend in changing terminology (Tantam, 2014).

For the purpose of this review, it is worth briefly mentioning the areas of the brain consistently associated with the ‘mentalizing network’ (Happé & Frith, 2013). The right and left temporo-parietal junction, precuneus, superior temporal sulcus and medial prefrontal cortex (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011 p705) are consistently activated by tasks that assess belief representation. Furthermore, atypical connectivity of this network is a hallmark of the biological basis for mentalizing difficulties observed in adults with ASD (Kana, Keller, Cherkassky, Minshew, & Just, 2009). Social cognition, by

virtue of its complexity, requires additional structures that are involved in emotion regulation, responsiveness and recognition as well as mentalizing (Kennedy & Adolphs, 2012). These are areas of importance when considering the wider concept of ToM and for further information see Happé and Frith (2013), Kennedy and Adolphs (2012) and Walter (2012).

1.3 Assessing/Measuring Theory of Mind

1.3.1 False Belief Tasks

The seminal work of Wimmer and Perner (1983) presented the ‘false belief task’ to measure ToM abilities in typically developing children. Within their paradigm, children are presented with a scene where a character (e.g. ‘Maxi’) leaves a desired item (e.g. some chocolate) in one location (e.g. the cupboard) and then leaves the scene. While he is away, Maxi’s mother comes into the scene and moves the chocolate. Upon Maxi’s return, the children are asked ‘Where will Maxi look for the chocolate?’ The authors argue that to pass, the children had to be able to represent a belief (Maxi *thinks* the chocolate is in the cupboard) that is distinct from their own regarding the situation, or *possess a first-order* ToM. They found that typically, around the age of 4, children passed this task and subsequent meta-analysis of false belief paradigms have revealed that age is a consistent predictor of task performance and differing results cannot be ascribed to ‘procedural differences’ e.g. whether a doll or a cartoon is used (Wellman, Cross, & Watson, 2001).

The paradigm was first used with young people with ASD by Baron-Cohen, Leslie and Frith (1985) who showed for the first time that 80% of children with ASD failed the ‘Sally-Anne’ task whereas 85% of neurotypical children (matched for ‘mental age’) and children with Down’s syndrome (matched for chronological age) passed. This result suggests that young people with ASD were unable to represent the mental states of other agents or lacked a ToM (Baron-Cohen, et al., 1985) and it was hypothesised that this deficit explained the impairments in social behaviour and imagination, characteristic of ASD (Hill & Frith, 2003).

A wealth of literature considering the pre-cursors to ToM development emphasising the roles of early social interactions such as joint attention and pretend play followed

(Baron-Cohen, 1989b; Charman et al., 1997; Leslie & Happé, 1989) . Theoretically, the development of such behaviours (or lack thereof in ASD) was considered evidence for a 'meta-representation system' account, necessary for later ToM development (Leslie, 1987; Leslie & Frith, 1988). While this theory has been well supported, it has not gone unquestioned (Leekam & Perner, 1991; Mundy & Sigman, 1989). Criticisms focus on whether the 'Sally Anne task' solely taps belief representation and inferring from this that it is a modularized cognitive system and, secondly, whether these abilities emerge around four years in typically developing children (Bloom & German, 2000). Firstly, the consistent association between domain general process (e.g. executive functions) with ToM development (Pellicano, 2007, 2010, 2012) questions the modularity assumption. Secondly, whether the false belief task itself is recruiting other cognitive processes has been debated (for more information on this debate see Bloom & German, 2000; Wimmer & Weichbold, 1994). Finally, the findings that 15 month olds pass the paradigm if looking time is analysed rather than verbal output (Onishi & Baillargeon, 2005), has led theorists to question when meta-representation emerges.

A key point that arises from research into false belief tasks is that for ToM abilities to be accurately assessed at different ages and abilities, one needs to apply a developmentally sensitive measure (Baron-Cohen, 2000). The means by which ToM is assessed is therefore of paramount importance in understanding atypical social functioning. The pursuit of such methodologies are equally important when considering early detection of precursors to later ToM difficulties in infancy (Jones, Carr, & Klin, 2008; Zwaigenbaum et al., 2009) as when assessing subtle ToM difficulties in adulthood (Roeyers & Demurie, 2010).

Results from early false belief studies showed that although a high percentage of individuals with ASD failed the tasks (40-80%; Happé, 1995), the fact that some passed suggested that having autism was not synonymous with first-order ToM impairment (Happé, 1994). Baron-Cohen (1989a) investigated the performance of young people who could pass the original first-order false belief paradigm (Baron-Cohen, et al., 1985) on a more sophisticated, second-order false belief task. The target question required an understanding of what 'Mary thinks that John thinks' (Baron-Cohen, 1989a) following a story where John's belief has been updated unbeknownst to Mary (e.g. Mary holds a false belief about John's belief). The entire autistic group failed on the task compared to just 10% of the neurotypical control group and 40% of the young people with Down's

syndrome. However, Bowler (1992) suggested that successful performance on second order ToM is possible by adults diagnosed with Asperger's syndrome whose verbal abilities are within the average range.

Compiling research investigating individuals with ASD performance on ToM tasks, Happé (1995) showed that individuals with ASD require a higher level of verbal abilities to pass false belief paradigms compared to typically developing youth. While the author was cautious not to ascribe a causal interpretation to correlation analysis, she suggested that individuals with ASD may be able to use their verbal abilities to 'solve' the false belief problem (Happé, 1995). The interpretation is supported by individuals with autism being able to explain their correct performance on second-order ToM tasks while younger typically developing 'passers' were not (Happé, 1995). Additionally, this points to an inherent association between linguistic performance and ToM, in that, to understand what is often being *communicated* by a speaker, a dynamic interplay between these two constructs is required (discussed in more detail below; Happé, 1993; Milligan, Astington, & Dack, 2007). These points are of particular importance to the current study as they suggest that instances where language is used in a pragmatic sense will cause individuals with ASD unique difficulties, but that verbal abilities are likely to play a role in performance (Happé, 1994; Sarah White, Hill, Happé, & Frith, 2009).

1.3.2 Overcoming Initial Problems with False Belief Tasks

A fundamental flaw in early ToM tasks, which partially undermined the ToM hypothesis, was the intact ability of older, high-functioning individuals with autism who performed akin to controls (Bowler, 1992). These findings called for the development of more advanced ToM measures to capture the subtle difficulties of this able group. Overcoming ceiling effects on traditional false belief tasks was first evidenced by Happé (1994) Strange Stories (SS) paradigm. The measure was rooted in 'relevance theory' which positions inference at the heart of human communication (Sperber & Wilson, 2002).

The decoding of verbal output is merely the first step to understanding, as it can often be undermined as in the case of irony. Meaning can therefore only be derived from a constellation of the verbal output alongside ostensive cues such as the speaker's actions, gestures and intonation or the context of the social exchange (Sperber & Wilson, 2002).

For Sperber and Wilson (2002) it is the *relevance* of either external (e.g. gesture) or internal (e.g. memory) inputs that guide the individual's inference of a particular meaning from the multitude possible in most social scenarios. The SS (Happé, 1994) used relevance theory to show comprehension of increasingly sophisticated linguistic concepts (simile; metaphor; irony) was associated with increasing ToM abilities because of the differing levels of inference needed to comprehend the intended meaning within each concept (Happé, 1993, 1995).

The SS consists of 24 short vignettes that include two examples of 12 forms of figurative language: Idiom, mixed emotions, misunderstanding, deception, forgetting, appearance reality distinctions, verbal irony, persuasion, joke, double-bluff, pretence and white lies. Control vignettes that required non-mental state reasoning to answer correctly were included to enable domain general processes to be accounted for and isolate unique difficulties with linguistic concepts that required mental state inference. Happé (1994) found that autistic individuals' performance on the SS was related to performance on traditional first-order and second-order false belief tasks. However, even the most intellectually able individuals with ASD were outperformed on the SS by all other control groups. When stratifying the results by incorrect responses, but still including reference to mental states (e.g. thoughts and feelings), the able ASD group responded with significantly more incorrect mental state answers. Shorter paradigms have included fewer stories and developed continuous rating systems (Fletcher et al., 1995; White, et al., 2009), which encourages use in a wider range of settings (e.g. clinical assessments).

A recent meta-analysis suggests that the SS are useful in differentiating ToM abilities in adults with HFA along with the 'faux pas test', a vignette-based measure requiring participants to judge whether a social faux pas has been committed (Chung, Barch, & Strube, 2013). The SS are sensitive to developmental changes in ToM across middle and late childhood (Devine & Hughes, 2013), and have been shown to be related to reading comprehension (Ricketts, Jones, Happé, & Charman, 2013). IQ is significantly associated with performance in a number of studies (Dziobek, et al., 2006; White, et al., 2009)

The SS are not without limitation however. The vignette design of the task undermines its utility, as the assessment of real social stimuli (e.g. facial expression, intonation) is not demanded. Furthermore, participants are given as much time as is necessary to process the material, making it deviate further from the fast-paced nature of social interaction,

which may explain why individuals ‘pass’ yet still struggle in day-to-day scenarios (Scheeren, de Rosnay, Koot, & Begeer, 2013). A number of studies were also not included in the aforementioned meta-analysis that compared the performance of adults on the SS. Senju et al., (2009) found differential gaze-inferred predictions in the ASD group to a standard false belief task and White, et al. (2009) revealed greater response latencies in the ASD group. This provides evidence for the notion that adults with HFS/AS may use compensatory strategies or alternated cognitive routes to ‘pass’ ToM tasks in the laboratory that would differentially affect real time social understanding (Bowler, 1992 208).

The SS has been the platform for conceptually similar measures that include dynamic material. One recent example is the ‘Silent Films task’ that took clips from a 1920’s silent film (Devine & Hughes, 2013), using incidences where the character’s behaviour could be best understood with mental state attribution e.g. misunderstanding. The measure is clinically useful as there are only 5 clips and each is short ($M=25$ secs). Convergent validity was established as performance correlated with the SS task, and inter-rater reliability was good (Devine & Hughes, 2013). The large sample enabled performance on the measure to be mapped onto a single ToM latent variable confirmed by factor analysis (Devine & Hughes, 2013). The nature of the stimuli is problematic however, for high functioning adults with ASD that have subtle impairments.

Firstly, the emotional expressivity in Devine and Hughes (2013) task is exaggerated as one would expect from a 1920’s silent comedy. The slapstick quality makes the ecological validity for the measure questionable. While it is an old film and the sample young, prior exposure to the material in an adult sample (e.g. the plot and/or storyline) could affect comprehension. The range of mental state concepts is also limited by its brevity. While the absence of verbal information make the inferences of the character’s behaviour reliant on integrating context, facial expression and body language the absence of control clips makes falsifying central coherence (Frith & Happé, 1994) as an explanatory variable in task performance challenging. The conceptual framework is useful however, as the measure is specific and focused on beliefs. The clips are short enough to place minimal demands on working memory and do not require updating a narrative or plot to understand the mental states being presented. Such advantages, make it a purer ToM measure than is often seen when using dynamic stimuli (Heavey, et al., 2000).

The 'Frith-Happé' triangles (F-HT; Abell, Happé, & Frith, 2000) is another non-verbal paradigm, in which participants view two animated triangles interacting with one another. The triangles' interactions were designed to be best described as random, goal directed (e.g. following) or intentional (e.g. persuading) and performance on the task was first shown to differentiate children with and without ASD (Abell, et al., 2000). Castelli, Frith, Happé, and Frith (2002) showed the animations to 10 adults with HFASD and compared their responses to matched controls finding that adults with HFASD were only differentiated on the intention animations. Individuals with HFASD were both less accurate and less likely to use psychological state talk (PST) when describing the animations' interactions. The findings evidenced the imbuing of geometric shapes by controls with intentions, and activation of the 'mentalizing' network despite the absence of any 'sentient' interactions. Conversely, this finding was not observed in the HFASD group. Finally, the reduced use of PST in communicating about interactions is of particular relevance to the conceptualization of HFASD and separates this group from controls. The short and applicable nature of the test makes it a useful measure however, the lack of ecological validity undermines its utility in the pursuit of an appropriate measure of subtle ToM impairments in HFASD (Dziobek, 2012).

1.3.3 Mentalizing and Emotion Recognition: More Ecologically Valid?

A range of mentalizing tasks were developed to bridge the gap between the laboratory and the social world that primarily focused on emotion recognition, including assessing such inferences from 'vocalizations' (Rutherford, Baron-Cohen, & Wheelwright, 2002) or from photographs of the eye regions alone e.g. the Reading the Mind in the Eyes Task (Eyes; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). The latter has been shown to be effective overall at identifying mentalizing difficulties in adults (Chung, et al., 2013). However, a recent critical review of the measure has suggested that the task's validity as a measure of mentalizing is undermined by the choice of stimuli (actual mental states of the images are unknown and were decided upon via consensus) and the distracter questions enable an alternate route to performance e.g. by exclusion (Johnston, Miles, & McKinlay, 2008). Furthermore, the role of VIQ has seldom been analysed yet a recent meta-analysis suggests that it plays a significant role in performance (Baker, Peterson, Pulos, & Kirkland, 2014)

Whilst using dynamic stimuli may make emotion recognition more ecologically valid it does not necessarily make it more challenging for those with ASD (Back, Ropar, & Mitchell, 2007). Of interest for the current study, emotion recognition was differentially harder for young people with ASD when only visual or prosody derived cues were available compared to when verbal or a combination of all forms was presented (Lindner & Rosén, 2006). These findings suggest a preference for verbal ostensive clues when decoding emotional stimuli.

Interestingly, no interaction effects between the HFASD individuals and controls were observed when ‘neutralising’ the eye/mouth region of images depicting a range of facial emotional expressions (Back, et al., 2007). Thus, while the ASD groups’ performance was poorer overall, removing information from the eye region or the mouth region affected both groups’ performance (Back, et al., 2007) suggesting that non-verbal ostensive cues also aid individuals with ASD’s performance. This group difference disappeared when the eyes were viewed in isolation (Back, et al., 2007), contradicting the findings of Baron-Cohen, Wheelwright, Hill, et al. (2001). Finally, the largest study to date revealed that on a test battery of visual and auditory (both verbal and non-verbal) emotion recognition tasks no differences were found between ASD and control groups and the pattern of errors did not differ (Jones et al., 2011). These mixed findings highlight that the availability of both non-verbal and verbal ostensive cues affect emotion recognition for individuals with and without ASD and that the role of VIQ may be of particular importance even in seemingly non-verbal tasks, suggesting again its close relationship with ToM (Milligan, et al., 2007). It may be the identification of complex emotions that are more challenging for those with HFASD (Golan, Baron-Cohen, Hill, & Golan, 2006). Alternatively, the emotions displayed in these paradigms were congruent with the ostensive cues available and thus situations where these contradict one another (as in sarcasm) may lead to particular challenges for individuals with ASD.

1.4 Ecologically Valid Tasks of Social Cognition.

The problem of assessing subtle ToM difficulties in an ecologically valid manner has led to a conceptual divide as to whether one should prioritise limited, but natural interactions (Roeyers, Buysse, Ponnet, & Pichal, 2001) or theoretically tailored, but acted stimuli (Dziobek, et al., 2006). The following sections will endeavour to cover this divide and highlight the strengths and weakness of both approaches.

1.4.1 Empathic Accuracy

The empathic accuracy paradigm was developed from the social psychology literature and could be argued to be the most naturalistic form of ToM task to date (Roeyers & Demurie, 2010). The paradigm was developed by Ickes, Stinson, Bissonnette & Garcia (1990) and involves secretly filming naive participants [targets] having a conversation while waiting to take part in an expected experiment. The targets then watch their conversation back and write down their thoughts and feelings at specific moments in the conversation. The term ‘empathic accuracy’ is the discrepancy between the target’s retrospective mental states and the study participant’s inferences of them when watching the conversation at a later point (Ickes, 1993).

Roeyers, Buysse, Ponnet, & Pichal (2001) first applied the paradigm to individuals with pervasive developmental disorder and then specifically individuals with AS (Ponnet, Roeyers, Buysse, De Clercq, & Van Der Heyden, 2004) and revealed that the ASD participants only showed difficulties relative to the controls in unstructured conversations. Thus, when the next logical path of a conversation can be expected or predicted, the mental states associated with it can be more readily inferred. For HFASD individuals, positive and neutral valence was beneficial to performance in more structured social exchanges (Ponnet, Buysse, Roeyers, & Clercq, 2008). An important finding for the current study was that when the targets’ internal states were incongruous with observable cues, these were the hardest to decipher (Ponnet, et al., 2008). The original paradigm has also been used with adolescents with ASD replicating the findings (Demurie, De Corel, & Roeyers, 2011).

The empathic accuracy paradigm has enabled the naturalistic assessment of social interaction that has demonstrated good reliability and validity (Marangoni, Garcia, Ickes, & Teng, 1995). However, the nature of a brief conversation with a stranger is less likely to be associated with particular linguistic constructs that people with ASD may struggle with (e.g. deception; Happé, 1994) especially as incongruence between the content of the internal state and the observable cues predicts poor performance (Ponnet, et al., 2008). The generic nature of the situation may also lead to a narrowing of potential internal states being experienced. Furthermore, the conversational content can only be as

sophisticated as one member of the partnership's social skills permit (Ponnet, Buysse, Roeyers, & Corte, 2005). For these reasons, while the empathic accuracy paradigm is heralded as the most naturalistic of social cognition tasks it is limited as a clinical tool by this strength.

1.4.2 Social Cognition Measures using Dynamic Stimuli

A number of studies have attempted to capture social cognition difficulties using dynamic stimuli. Table 1 below shows current tasks of social cognition that have utilised dynamic stimuli to assess social cognition abilities in the HFASD populations.

Table 1: Characteristics of current dynamic social cognition tasks.

| Author | Test | Stimuli | Question type | Participants | Relevant findings | Strengths | Limitations |
|------------------------|------|--|--|-------------------------------------|---|---|--|
| Heavey et al., (2000) | AMT | UK advertisements (7) and TV series clip (1). | FC ER FC memory Open ended interview regarding intentions of characters. | Adults: 16 HFASD 15 Controls* | HFASD < Controls, including some Memory questions. Intention yielded greater effects than FC ER questions. Only controls performance on AMT related to the SS and IQ. No group response latency difference | Open ended questions. Convergent Validity | HFASD group struggled with memory questions. Complex coding system for intentionality. 45-120 second long clips. Overacted/dramatic stimuli. No control clips. |
| Golan et al., (2006) | RMFT | 22 short film clips from feature films. | FC ER | Adults: 22 HFASD 22 Controls* | HFASD < Controls Performance on RMFT related to VIQ, AQ and CMFVB | Replicated with child version Complex emotions. Convergent Validity | No control clips/questions. Consensus decided emotions. |
| Dziobek et al., (2006) | MASC | 15min video of 4 characters preparing for a party. Film stopped for each question(46 times). | Open ended concerning characters' thoughts, feelings and intentions. Memory. | Adults: 19 HFASD 20 Controls* | MASC group difference > Eyes, SS and ER task. HFASD=Controls on Memory Questions. No association with MASC and VIQ MASC associated with SS and ADI-R No association between Eyes, ER or SS tasks. | Open questions Tailored stimuli Range of linguistic concepts Convergent validity Re-test reliability Replicated with FC version. | 45min administration time. Non-English speaking. Trained rater required for scoring Basic control questions. |

Table 1: *Age, gender and IQ matched AMT, Awkward Moments Test, RMFT, Reading the Mind in the Films Task, MASC, A Movie for the Assessment of Social Cognition, FC, Forced choice, ER, Emotion Recognition.

Table 1 cont: Characteristics of current dynamic social cognition tasks.

| Author | Test | Stimuli | Scoring/ Question type | Participants | Relevant findings | Strengths | Limitations |
|-------------------------------|----------------------|---|--|---------------------------------------|--|---|---|
| Beaumont and Sofronoff (2008) | ATOMIC | 18 computer generated animations. | FC ToM, CC and memory. | Children: 25 HFASD 25 Controls* | HFASD < Controls, but not on CC or Memory questions. No difference between affective (feeling) and cognitive (thinking) ToM questions. Positive association between SS and ATOMIC No group difference in response latency | Differences independent of memory. 25min total administration time. Convergent Validity | Cartoon stimuli No control clips Child sample Limited diagnostic information. |
| Barnes et al., (2009) | MDFT | 4 film clips from American TV show 'House'. | Use of mental state words in narrative description of task, length of description, type of mental states used. | Adults: 28 HFASD 28 Controls* | Lower frequency of mental state references in HFASD narratives and shorter overall. VIQ correlated with performance only for HFASD. Empathy scores correlated with only controls' performance on MDFT. | Open questions. Convergent validity. | No intention questions. Dramatised stimuli. |
| Mathersul et al., (2013) | TASIT: part 2 and 3. | 31 self-contained clips of ambiguous social interchanges. | FC regarding thoughts, feelings (ER) and intentions of characters. | Adults: 40 HFASD 37 Controls* | HFASD < Controls, but not on ER questions. VIQ did not correlate with performance on TASIT. Only self-reported cognitive empathy predicted by TASIT independent of group. | Large sample. Convergent validity Bespoke clips | No control clips or questions. Lengthy administration (60-75mins). Unnatural acting** |

Table 1. ** Opinion of the author. *Age, gender and IQ matched. ATOMIC, The Animated Theory of Mind Inventory for Children, MDFT, Moral Dilemmas Film Task, TASIT, The Awareness of Social Inference Test, FC, Forced choice, ER . Emotion Recognition.

A number of conclusions can be drawn from reviewing the findings of the studies in Table 1. Firstly the studies consistently found mentalizing difficulties in the HFASD group compared to well matched controls (Barnes, Lombardo, Wheelwright, & Baron-Cohen, 2009; Beaumont & Sofronoff, 2008; Dziobek, et al., 2006; Golan, et al., 2006; Heavey, et al., 2000; Mathersul, McDonald, & Rushby, 2013b). The convergent validity of the new measures were supported in most cases and a number of the tasks have been either replicated with other samples (Golan, Baron-Cohen, & Golan, 2008; Montag et al., 2011; Sparks, McDonald, Lino, O'Donnelle, & Green, 2010) or had meaningful forced choice versions created (Montag, et al., 2011). The majority of the studies employed forced choice questions. One concern with such designs is that orientating participants to consider feelings, thoughts intentions may differentially affect the degree of mental state language produced as compared to being unprompted (Callenmark, Kjellin, Rönqvist, & Bölte, 2013) and the deviation from day-to-day interchanges such directive probing creates.

Administration time and scoring complexity are evident weaknesses however, to the studies that employed open-ended questions. VIQ did not undermine group differences on the social cognition tests, however, the studies yielded differing levels of associations between performance and verbal IQ within groups (see Table 1) questioning the necessary association of the two constructs. In the studies that did assess response latency, no differences were observed (Beaumont & Sofronoff, 2008; Heavey, et al., 2000) suggesting that the HFASD individuals may be processing the material as automatically as the controls.

A consistent limitation with all of the studies, was the lack of well-designed control items that matched the experimental stimuli in cognitive difficulty. Memory questions at times were employed but often reached ceiling (Dziobek, et al., 2006) and in some cases yielded group differences (Heavey, et al., 2000) highlighting the necessary inclusion of such material to conclude *specific* mentalizing difficulties from group differences.

The majority of the studies also utilised pre-existing film material that is often highly dramatised or chosen because of affective content (Golan, et al., 2006). In such cases, defining the emotion displayed becomes a consensus decision and the naturalistic quality of the exchange is affected. The use of animated stimuli (Beaumont & Sofronoff, 2008) undermines ecological validity and relevance to real world social exchanges, which has been argued as essential to understanding the social abilities of more able individuals with

ASD (Dziobek, 2012). Finally, one of the more psychometrically and theoretically robust measures was filmed in a non-English speaking country, with the English version being dubbed (Dziobek, et al., 2006). To the authors' knowledge the dubbed version has not been validated with an English sample. Recent research from a Spanish (SP) dubbed version has yielded differing results from the original paper, with the Eyes test revealing greater group sensitivity than the MASC-SP (Lahera et al., 2014). With regard to the literature on preferential gaze in ASD samples to the mouth region (Ami Klin, Jones, Schultz, & Volkmar, 2003), dubbing stimuli may differentially affect ASD individuals when inferring mental states from such material.

Overall, a range of theoretically devised and well validated measures are currently available to assess mentalizing in HFASD. The limitations in appropriate control material, dramatised stimuli and designs limited to 'short but forced choice questions' and 'lengthy, but open-ended questions' suggests this area still requires development.

1.5 Cognitive vs. Affective Empathy

Empathy is a complex construct to define with no universally agreed definition, but is argued to be a central feature of ASD (Baron-Cohen & Wheelwright, 2004). A helpful heuristic in the empathy literature is dividing the construct into two psychological processes; cognitive and affective empathy (Davis, 1980; Lockwood, Bird, Bridge, & Viding, 2013). The former is associated with abilities in representing the mental states of other people and in turn predicting their behaviour; such definitions share a clear similarity with ToM (Baron-Cohen, et al., 1985; Premack & Woodruff, 1978) and researchers use the terms interchangeably (Rogers, Dziobek, Hassenstab, Wolf, & Convit, 2007; Smith, 2009a). Affective empathy is an emotional state that is triggered in response to an observed or imagined emotional experience (de Vignemont & Singer, 2006).

The two subsystems are hypothesised to be 'related but independent' (Rogers, et al., 2007), thus it is possible to have an affective response without a cognitive component and vice versa (Smith, 2006, 2009a, 2009b). The fractionation or dissociation of these components of empathy has been particularly relevant to the ASD literature and in particular, dispelling the misperception that autistic individuals are generally lacking empathy (Huws & Jones, 2010; Smith, 2006). Furthermore, the interaction and biological

basis of these constructs has yielded a consistent pattern of brain regions that are associated, but distinct from each other (for a review see Walter, 2012) .

Rogers, Dziobek, Hassenstab, Wolf & Convit (2007) compared a group of adults with a diagnosis with ASD with matched neurotypical controls on the interpersonal reactivity index (IRI); a self-report measure that differentiates between the cognitive and affective dimensions of empathy (Davis, 1983). They found a preservation or heightening of the affective components of empathy, with 'personal distress (PD)' being significantly elevated in the autistic group and 'empathic concern (EC)' being intact. Conversely, the cognitive empathy dimensions; perspective taking (PT) and fantasising (F), were significantly lower in the ASD group. The profile of the ASD group suggested a greater sense of personal discomfort in difficult interpersonal situations, which is at odds with the lay conception of autistic individuals lacking empathy (Rogers, et al., 2007). The authors hypothesise that this heightened score may be related to difficulties in emotion regulation (Baron-Cohen & Wheelwright, 2004) although other theorists argue that it reflects a genuine hypersensitivity to emotionally laden stimuli (Smith, 2009b).

Interestingly, self-reported cognitive empathy was correlated within the control group to performance on the SS (Happé, 1994) suggesting that behavioural measures of ToM are associated with self-reported cognitive empathy. Mathersul (2013b) similarly found an association between questions on the TASIT pertaining to beliefs, intentions and meaning (ToM) and cognitive empathy as measured by a composite score of the IRI and Empathy Quotient (EQ). Affective empathy was only predicted by group (HFASD or control) despite an emotion recognition question being used as a predictor.

In contrast to previous literature, Mathersul et al., (2013b) findings that affective empathy was reduced in individuals with an ASD is contrary to research data showing that empathy is intact (Dziobek et al., 2008; Rogers, et al., 2007) and or heightened (Smith, 2009a). Importantly, Mathersul et al., (2013b) used a composite score from the EQ (Baron-Cohen & Wheelwright, 2004) and only the empathic concern subscale of the IRI. The authors argue that compared to the personal distress sub scale for the IRI the empathic concern scale is a purer measure of affective empathy (Mathersul, et al., 2013b). However, the researchers' differing concepts of what represents true affective empathy makes conclusion regarding whether it is reduced, intact or increased in ASD challenging (Smith, 2009a).

Demurie et al., (2011) found an overall dampening of both cognitive and affective empathy in a group of adolescents with ASD as measured by the IRI. However, they also found unique discrepancies between self-reported and parent-rated cognitive and affective empathy subscales. While it is difficult to ascribe the direction of discrepancy (e.g. are the parents or individuals less accurate), only the ASD group significantly over-reported their perspective taking abilities and under-reported their personal distress scores relative to parents. The literature on the reliability of self-report measures in adults with HFASD for emotional difficulties is mixed (Berthoz & Hill, 2005; Mazefsky, Kao, & Oswald, 2011), and discrepancy studies have mainly investigated adolescents (Storch et al., 2012).

The observed dissociation between the cognitive and affective components of empathy in both experimental tasks and self-report measures has also been reported in children and adolescents with ASD (Jones, Happé, Gilbert, Burnett, & Viding, 2010) and in community samples that score highly on ASD trait questionnaires (Lockwood, et al., 2013). It is this distinct empathy profile that researchers argue distinguishes autistic individuals from psychopathic individuals (another pathology characterized by disturbances in empathy), where the opposite profile is observed e.g. low affective empathy and high cognitive empathy (Blair et al., 1996; Blair, 2008; Dolan & Fullam, 2004).

Finally, this dissociation has been demonstrated in fMRI studies whereby components of the metalizing network showed reduced activation in autistic individuals when observing images of harm that was intentionally inflicted, coupled with heightened arousal (Fan, Chen, Chen, Decety, & Cheng, 2013). Furthermore, facial electromyography responses to emotional stimuli, which have been found to correlate with affective empathy (Sonnby-Borgström, 2002) have shown to be intact or heightened in autistic individuals (Magnée, De Gelder, Van Engeland, & Kemner, 2007; Mathersul, McDonald, & Rushby, 2013a) However, Mathersul et al., (2013a) showed that when measuring autonomic arousal via skin conductance and changes in heart rate autistic individuals did not follow a typical trajectory but instead were 'dampened overall' (Mathersul, et al., 2013a p20). The authors suggest one explanation could be a detaching from the affective stimuli to down regulate heightened arousal (Hadjikhani et al., 2014) as increased arousal has been shown when autistic individuals are exposed to social stimuli (Kylliäinen et al., 2012)

Overall, these findings suggest that ASD are characterised by reduced cognitive empathy profile, which supports the mentalizing/theory of mind literature. While not entirely consistent, the evidence suggests an association between some behavioural measures of ToM and the self-reported construct of cognitive empathy. Whether affective empathy is reduced or intact in autistic individuals is less consistent however at the physiological, cognitive, and behavioural level. That said, behavioural ToM measures appear to be consistently tapping a separate construct to self-reported affective empathy, as no previous research has shown an association between these two constructs.

An alternate hypothesis to account for the inconsistent findings in research assessing affective empathy and emotion processing differences is that a confounding variable is accounting for the variance, rather than participant's diagnostic category. Unmeasured alexithymia has been argued to potentially account for these observed differences (Bird & Cook, 2013) and is discussed in the next section.

1.6 The Alexithymia Hypothesis

Alexithymia is a trait defined as a person's inability to identify and describe their own emotions (Taylor, Michael Bagby, & Parker, 1991) and within the alexithymia construct there are both cognitive and affective components (Grynberg, Luminet, Corneille, Grèzes, & Berthoz, 2010). Alexithymia is common in a range of clinical groups, including those with eating disorders, psychosis and ASD. The suggestion is not that autism and alexithymia are a unitary construct, but that they overlap, thus the emotional difficulties observed in the ASD population can be explained by differences in alexithymia rather than ASD per se (Bird & Cook, 2013). Such a hypothesis would predict that individuals high in alexithymia traits would struggle in the areas of emotion recognition and empathy like those with ASD.

Moriguchi et al. (2006) found that those high in alexithymia traits showed lower PT, EC and higher PD scores on the IRI and performed poorly on the Frith-Happé triangles task. In addition, Swart, Kortekaas, and Aleman (2009) found that those high in alexithymia traits scored lower on the EQ and were less able to recognise subtle emotional expressions. Within ASD samples, higher incidences of alexithymia have also been observed (Hill, Berthoz, & Frith, 2004) adding to the evidence that these two constructs

are related. It is important to add however, that the scales used to measure alexithymia (through self-report) have shown differing abilities to differentiate ASD from control groups at multiple time points (Berthoz & Hill, 2005). In Berthoz and Hill's (2005) research the Toronto Alexithymia Scale (TAS-20; Bagby, Parker, & Taylor, 1994) was deemed the most reliable self-report measure of alexithymia traits in their ASD sample.

A number of research studies show alexithymia traits in those with autism being closely associated with differences in performance on both visual and auditory (both verbal and non-verbal) emotion recognition paradigms (Cook, Brewer, Shah, & Bird, 2013; Heaton et al., 2012) and a recent review supports these findings (Grynberg et al., 2012). Furthermore, once alexithymia was controlled for, group differences in activation of brain regions associated with observing others in pain did not remain significant (Bird et al., 2010). However, some caution should be exercised when interpreting the findings. Heaton, et al. (2012) only reported correlations between the TAS-20 and emotion recognition findings and concluded that alexithymia could not account for all the specific difficulties observed in the ASD group. Again, while Fan, et al. (2013) found a correlation between alexithymia and affective empathy, it was not a 'significant mediator in processing pain empathy' (Fan, et al., 2013 p21). Finally, in the typically developing literature, recognition of non-verbal affect from prosody was not shown to be impaired in those high in alexithymia traits (Swart, et al., 2009).

In participants with ASD both affective and cognitive subscales of the IRI have shown to be associated with alexithymia traits (Silani et al., 2008). In contrast, research using a large community sample showed alexithymia was not associated with difficulties in cognitive empathy observed in participants who showed high ASD traits, but was independently associated with affective empathy difficulties in those high in psychopathic traits (Lockwood, et al., 2013). Importantly, however, not all the variance in the latter finding could be explained by alexithymia (Lockwood, et al., 2013).

Overall, the research suggests that alexithymia is less associated with the cognitive elements of empathy/ToM. No research has investigated the relationship between more sophisticated and ecologically valid measures of social cognition and how alexithymia traits relate to performance on such measures.

1.7 Successful Mentalizing Equals Appropriate Social Interaction?

Even when people with HFASD can track others' mental states they may have difficulty generating appropriate behaviours in an interaction. There is relatively little research on this topic, but research on moral cognition in ASD may be of relevance.

A range of recent studies have begun to identify how HFASD individuals use their ToM abilities to make moral judgements (Zalla, Barlassina, Buon, & Leboyer, 2011), tailor their responses using mentalizing information, or respond to characters who might be in need (Jameel, Vyas, Bellesi, Roberts, & Channon, 2014; Yang & Baillargeon, 2013)..

The 'Above and Beyond (AaB)' task developed by Jameel, et al. (2014), is a vignette-based task describing characters in need e.g. a close friend breaking up with a partner. The researchers asked 'What would you do in this situation?' as a measure of unprompted pro-social behaviour. They found that the participants high in ASD traits were less pro-social overall. The concept of pro-social behaviour is linked to Yang and Baillargeon (2013) ideas of 'social acting', which is where beliefs are decoupled from a speech act for the benefit of another (e.g. white lies). In both, the perspective of the other is overriding the perspective of the individual, and the former guiding behaviour. Yang and Baillargeon (2013) found it was 'social acting' and not mentalizing that predicted quality of social relationships. These findings highlight a gap in the ToM literature, which is particularly pertinent to the current study. Until recently, the focus has been mainly on internal state representation. How such knowledge is applied however, is equally important, especially when findings have shown that mental state inference can be highly successful in adults with HFASD (Ponnet, et al., 2005) even though adaptive behaviour and abilities to maintain meaningful relationships can be limited (Palmen, Didden, & Lang, 2012).

Another area investigating the relationship between behavioural responses and ToM is the field of moral decision making. Moran et al. (2011) has shown that adults with HFASD rate characters' behaviour as less permissible only in vignettes that result in a negative outcome, but the behaviour is unintentional. Similarly, Buon et al. (2013) created silent cartoons where the characters' intentions, responsibility and causal role in a situation were manipulated. The HFASD participants were more likely to ascribe higher levels of intentionality, responsibility and punishment in scenarios where negative outcomes were accidental. While Buon, et al. (2013) argue that, in part, this finding is related to reduced mentalizing abilities, both Moran, et al. (2011) and Buon, et al. (2013)

research suggests that for the more able HFASD participants, information regarding the character's mental states is being undermined by outcome, which has also been shown in positive outcomes (Zalla & Leboyer, 2011).

In all, these findings are particularly interesting as they point to a decoupling of the ability to understand others' mental states and the ability to use this information to guide decision making and inform behaviour, which is not solely a bi-product of limited affective empathy (Gleichgerricht et al., 2013). Throughout all of this research, however, the use of paper tasks to measure mentalizing abilities in HFASD individuals suggests that their abilities may be being over estimated. Better measures of ToM would help tease apart those individuals who were unable to use this information to guide behaviour, even when passing more sophisticated measures of mentalizing.

1.8 Aims of the Current Study

The current study aims to develop and evaluate a new video-based mentalizing task that addresses a number of the gaps in the literature and can differentiate between adults with HFASD and matched controls. The Strange Stories Film Task (SSFt) will use dynamic stimuli, requiring real time processing of social exchanges with naturalistic emotional expressions in day-to-day scenarios, which are not exaggerated or unusual. The task will be designed to avoid the ceiling effects undermining healthy controls' performance. Furthermore, the measure will focus on figurative language to build on the vast literature suggesting that mentalizing difficulties are related to the degree of inference necessary to understand the communicated meaning of a speaker. The measure will include a behavioural response question to build on recent literature exploring the relationship between mental state inference and appropriate behavioural responses. To account for more domain general processes control questions and control clips will be included in which understanding of mental states is not necessary, but the appropriate integration of verbal/contextual information and recruitment of attention resources is required. To assess the validity of the new measure, traditional, well validated measures of ToM and emotion recognition will be included in the test battery. Similarly, self-report measures assessing potentially associated variables will be included (alexithymia and empathy) to further explore the relationship between these constructs in ASD and control populations.

1.9 Hypotheses

1.9.1 Group Differences

1. **The HFASD group will show lower scores on the experimental stimuli compared to the control group.**
2. **The HFASD group and the control group will perform equally on memory questions and all the control stimuli questions.**
3. **The HFASD group will show lower scores on the SS, Eyes, F-HT but not on the TASIT.**
4. **SSFt will be more effective at differentiating between the two groups than the SS, Eyes task, F-HT and the TASIT.**

1.9.2 Convergent Validity

5. **Performance on SSFt will be positively associated with the SS and the F-H T triangles ,TASIT and the Eyes tasks.**
6. **Performance on the SSFt will be positively associated with the PT and F subscale of the IRI, but not the affective subscales.**
7. **Performance on the SSFt will be negatively associated with the TAS-20.**
8. **Performance on SSFt will be correlated with reciprocal social interaction scores and communication scores on the Autism Diagnostic Interview-Revised in the ASD group, and the AQ in the controls.**
9. **Performance on the SS, Eyes, F-HT but not the TASIT will be associated with the ADI-R communication and social domains and the AQ.**
10. **Performance on the TASIT and the Eyes will be positively associated**
11. **Performance on the F-HT and the SS will be positively associated.**

1.9.3 Empathy

12. **The HFASD group will show lower scores on the cognitive domains of the IRI but not on the affective subscales.**
13. **Performance on the SS and F-H T will be positively associated with the cognitive domains of the IRI but not the affective subscales.**

1.9.4 Alexithymia

14. The HFASD group will report higher levels of alexithymia than the controls.

15. Performance on the Eyes and the TASIT will be negatively associated with the TAS-20.

2 Method

2.1 Devising the Strange Stories Film Task (SSFt)

2.1.1 Scripts

Scripts were informed by clinical experience, research literature and service user consultation. A copy of the scripts can be found in Appendix 1. Prior to embarking on this project the author and second supervisor worked in an adult autism service, both providing psychological therapy and completing neuropsychological assessments. A repeated theme from service users struggling with social situations was surrounding language that was both non-literal and subtle; for example, when people are sarcastic, but do not exaggerate the fact. One service user reported his partner coming home (where he had spent the day working) and his belongings were everywhere. His partner stated “this is just perfect” so he left his items on the table, misunderstanding the intention behind her statement. This situation described by the service user was used as a basis for one of the verbal irony scenes.

The SS devised by Happé (1994) were used as a platform to inform the types of situations and communication that individuals with ASD are likely to find challenging. It provided a guide to the overall structure of the assessment tool. It was the authors’ intention to avoid using clips from films (Devine & Hughes, 2013; Golan, et al., 2006; Heavey, et al., 2000) to control for the effects of prior exposure to the material, overdramatized expressions, distracting camera work (e.g. fast paced editing) and unrealistic scenarios that are unlikely to occur in day-to-day interactions. To reduce cognitive load, memory and attention, the scenarios and accompanying dialogue were intentionally kept short with the majority of scenes only consisting of a single utterance per character. Participant’s did not have to follow a narrative, plot or back story to answer the questions as is necessary in the MASC (Dziobek, et al., 2006) as each script was self-contained.

The language used in the scripts was kept as close to everyday spoken language as possible and complex or overly sophisticated vocabulary was avoided. When filming took place it was not always possible to adhere rigidly to the original script and the actors were

given artistic license to fine tune subtleties as long as the themes, affect and intended meanings were maintained. Similarly, at times body language, gestures and expressions were included that were not scripted to keep the clips from appearing 'wooden'. The scripts were written to include day-to-day scenarios that adults were likely to be faced with in their lives. Scripts were intentionally written with different emotional valence including both positive and negative affect. Three or four scripts for each theme present in Happé (1994) SS were written to enable sub-optimal clips to be deleted from the final version.

In addition, ten control scripts were written (for a copy see Appendix 2). These mirrored the experimental clips in terms of length, cognitive load and linguistic sophistication. However, they required logical reasoning, (e.g. economic decision making or natural phenomenon), to decipher the characters' language or behaviour, rather than understanding of mental states per se, akin to the control vignettes used by Fletcher, et al. (1995) and White et al (2009).

2.1.2 Actors

An advertisement was placed on www.castingcallpro.com with an accompanying brief for the project. 83 applicants responded to the advert. From this 17 were deemed suitable and could make the 2 day audition period. Each was sent a brief of the characters and the project (see Appendix 3). The potential actors performed a range of scenes with the lead author and producer playing the other character. Each scene was filmed to camera. Of the 8 applicants who attended audition, the final two were chosen based on their portfolio (which included a lead character in a television series for the lead female actor) ability to respond to direction, naturalistic performance and the plausibility of the scripted relationship e.g. age and appearance. The actors were not paid and took part in the project both to enhance their show reel and because they deemed the project worthwhile.

2.1.3 Filming and Editing

Filming was completed over a three day period that included six locations; a residential property, local pub, green grocer, park, local street and office space. The

majority of the scenes were filmed within the residential property. In each scene, a third person perspective shot was filmed to show the viewer the context of the social exchange. The scenes of this initial shot were kept as sparse as possible e.g. artwork was taken from the walls, to minimise possible distractions that might differentially distract individuals with ASD (Klin, et al., 2003), but were still kept naturalistic (e.g. scenes were identifiable as occurring within a kitchen). The actors wore clothing that did not have any obvious logos or distracting patterns, a plain jacket and plain smart office clothing. During the indoor scenes outfits were rotated to convey that each scene occurred at a different time and were not part of a continuous narrative. In all office-based scenes the actors wore smart clothing and when outside a winter jacket. The female actor kept her hair tied back in all but one scene (a bedroom scene) so that her hair was not masking her face. All speech was directed to camera and thus filmed in the first person to both reduce possible attention biases for the viewers with an ASD (Klin, et al., 2003) and simulate a live conversation for the viewer as closely as possible. (see section 2.1.4 for screen shots of an example scene). A number of takes were performed until the intended balance of natural performance, emotional valence and smooth cinematography was achieved.

The clips were all filmed on a Sony Z4 camera and edited on Final Cut Pro on a Macintosh Laptop. During the editing process, takes that conveyed the closest depiction of the scripted affect and meaning were chosen through consensus between the author, producer and supervisor.

2.1.4 Questions

A series of questions were used to assess social understanding in participants immediately following their viewing of each clip. See example script, accompanying screen shots and questions below:

White lie example clip:

Third person perspective of Max and Alice sitting in the living room across from each other and Alice holding a guitar about to play:



Focus on Alice from Max's perspective: (looking nervous) 'I've been working on this for ages and I think I have finally got it. I think my songs gonna end like this.... (strums badly played chord then sings out of tune) ooo ooo ooo yeah' (looks expectant at camera)



Focus on Max from Alice's perspective: (nods head encouragingly and half smiles)



'Well done Alice... that sounds really good'

1. **Intention Question:** Why did Max say that?
2. **Interaction Question:** If you were in Alice's situation, what would you say next?
3. **Memory Question:** What instrument was Alice playing?

Due to the theoretical and structural similarities, the first question used in the measure was taken from Happé's (1994) SS 'Why did X [the character] say that?', and always referred to the last speaker and utterance. This wording of the question is sufficiently open ended, but also is directive enough for participants to consider the mental states of the characters. The statement to which this question refers will be referred to as the target utterance hereafter.

The second question was introduced based on clinical experience and consultation with an individual with ASD. The second question was: 'If you were in Y's [other character i.e. not X/the last speaker] situation, what would you say next?' This question was again deemed sufficiently open ended to gain a rich description from the participants. Moreover, it also targeted a common difficulty that was expressed by a number of clients in the aforementioned clinic and the service user consultant with a diagnosis of ASD; the ability to understand another's intentions without being able to generate a response that incorporates the mental states and intention of the speaker and thus continue the social exchange. These difficulties are highlighted in well validated measures of ASD such as question 26: 'I frequently find that I don't know how to keep a conversation going', in the Autism Quotient questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001a).

The final question was used as a means of controlling for potential lapses in attention or gross difficulties in memory abilities that may affect performance and took the form of a closed question pertaining to a piece of information in the scene, which was important for an appreciation of the scene's context and/or the interaction.

2.1.5 Order

The order which the items were presented was quasi-randomised. Initially, all the clips were assigned a random number using a random number generator website <http://www.pangloss.com/seidel/rnumber.cgi>. The 10 control clips were randomly

interspersed with experimental clips. This preliminary sequence was checked and re-arranged according to the following criteria:

- 1) **The first clip could not contain scenes where the relationship was not clear or a context had not been previously established, for example, one clip in which the characters are cross dressing.**
- 2) **Clips of the same theme (including control clips as a theme) could not directly follow one another, to avoid potential contamination or priming effects that would bias responses for that theme.**
- 3) **Clips that were similar in content, but crossed themes were not presented in succession. This was to reduce the likelihood that the participants' responses would have been affected if they could not consider each scene to be self-contained.**

Once the clip order did not violate criteria 1-3 it was labelled order A, which was then reversed and the same criteria outlined above were applied, to form order B.

2.2 Piloting ---

2.2.1 Ethics ---

Ethical approval was granted by the King's College London Psychiatry, Nursing and Midwifery Ethics Sub-Committee PNM/10/11-22. Information sheets were provided to inform potential participants of the nature of the study and written consent was gained prior to taking part in the study. Both these forms can be found in Appendix 4.

2.2.2 Participants ---

20 healthy adults took part in this part of the study and were recruited via an opportunity sample. The sample included 10 females and 10 males and the mean age of the samples was 28.8years (SD = 7.66). Participants were only recruited into the study if they had an Autism Quotient score below 32, which is the cut off suggestive of ASD/ high ASD traits (Baron-Cohen, Wheelwright, Skinner, et al., 2001a). No participants who opted

into the study had to be rejected from the pilot due to the presence of ASD traits as measured by the AQ ($M = 10.68$ $SD = 3.87$ range = 6-17).

2.2.3 Procedure

Participants completed the AQ and then returned the completed questionnaire. Participants then completed the SSFT Pilot on a laptop (using headphones to listen to the materials), within a quiet room at the Institute of Psychiatry. Participants answered the questions verbally and their responses were recorded verbatim on a laptop computer.

2.3 Measures

2.3.1 The Autism Quotient (AQ)

The AQ (Baron-Cohen, Wheelwright, Skinner, et al., 2001a) is a 50 item self-report questionnaire designed to assess ASD traits in adults with an Intelligence Quotient (IQ) in the normal range. A copy can be found in Appendix 5. The 50 questions are equally divided across five areas; 'social skill, attention switching, attention to detail, communication and imagination' (Baron-Cohen, Wheelwright, Skinner, et al., 2001 a, p.6), assessing both symptom clusters and well evidenced cognitive styles associated with the diagnosis. Scores of 0-1 are given for each question on a Likert scale; 'definitely agree, slightly agree, definitely agree, definitely disagree', where either the agree statements or disagree statements are awarded a point. Total scores range therefore from 0-50. Participants are asked to complete the questionnaire as quickly as possible, but there are no time limitations.

The AQ shows good test-retest reliability ($r = 0.7$) and the cut-off of >32 has been shown to discriminate adults with an ASD diagnosis from control groups ($N=1014$; Baron-Cohen, Wheelwright, Skinner, et al., 2001a). The questionnaire has been translated into 25 different languages and has been validated cross culturally, showing satisfactory internal consistency (Hoekstra, Bartels, Cath, & Boomsma, 2008; Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006).

2.3.2 The Strange Stories Film Task Pilot

The SSFt-p set consisted of 48 clips. Thirty-eight clips followed the themes of the 12 types of mental state vignettes presented in Happé's (1994) *Strange Stories*; lie, irony, double-bluff, pretence, joke, appearance reality, white-lie, persuasion, misunderstanding, forget, contrary emotions and idiom. Ten clips, which acted as control stimuli, were based on White, et al., (2009) physical state reasoning stories. The order of viewing was quasi-randomized as per the procedure outlined in section 2.1.5, and half the participants (stratified by gender) viewed order A and half viewed order B.

Prior to the task the experimenter read a set of instructions informing the participants of the nature of the task, the characters' relationship and invited any further questions (see Appendix 6). Each of the clips was approximately the same length and no clip exceeded 36 seconds ($M = 17.92$ seconds, $SD = 6.34$). Before each scene a black background was presented for one second and then the clip's number was presented on a black background for two seconds. The full running time of the pilot was 17 minutes and eight seconds.

After viewing the first clip the video was paused and the experimenter asked the three experimental questions outlined above in section 2.1.4.

The participants' responses to each question were typed verbatim on a scoring sheet (see Appendix 7). The next clip was then viewed and the procedure repeated. Following the end of the scenes, participants were thanked for their time and asked if they had any questions.

2.4 Selection of Final Test Set

2.4.1 Scene Selection

Appropriate scene selection for the final 12 stimuli followed a step by step procedure in order of priority:

- 1) **Ineffective clips where viewers did not identify the exact intended meaning (<5/20) in answering the first question were removed (6 experimental and 2 control scenes), if the clip did not match the intended meaning (1 experimental), if the theme was not equally portrayed (n=1) or the introduction of a new character (n=1)**
- 2) **Scenes were selected based on who delivered the target utterance (male or female actor), and setting (kitchen, living room, outside, in an office) with the aim of having a balanced set of scenes.**

The final set consisted of 12 experimental (one of each theme) and 3 control clips, where the female actor delivered the target utterance on nine occasions and the male on six. The scenes selected were shot within a living room on six occasions, a kitchen on six and outside/in an office on three. A second set of 12 viable clips was left for future research purposes (See Appendix 7 for scene selection).

2.4.2 Order effects

One issue that arose from the pilot was that participants had to adjust to the task and found answering the questions for the first time difficult. This was evidenced by comparing responses for item 1 in order A to item 48 in order B (the same clip) and vice versa. This prompted the implementation of three practice clips at the start of the final version of the task; two with social content and one control. This enabled participants to adjust to the task demands before starting the scored part of the task, and allowed the researcher to check whether the participant had any questions regarding the task before starting the experimental stimuli.

2.4.3 Imposing a Narrative

On a number of occasions the participants made reference to what happened between clips or referenced behaviour in one clip to guide their response in later clips, despite the explicit instruction to treat each as a self-contained clip. Following the pilot, prompts were provided by the experimenter during testing (for example 'remember that each story is self contained') if participants responded in a manner that suggested they were

connecting clips, e.g. explicitly mentioning information from previous clips when answering latter clips.

2.4.4 Prompting

For question 2 participants sometimes responded with a vague or generic answer such as ‘I would say something in response to the joke’. In these instances a prompt was provided (‘For example?’ or ‘Such as?’) to the participant to generate a more specific answer.

2.5 Scoring

2.5.1 Experimental Clips

The scoring system for the new task followed that described in White et al., (2009), Devine and Hughes (2013) and Castelli, Frith, Happé and Frith (2002). Of particular importance, this system has shown to be reliable in film-based tasks based on the SS (Devine & Hughes, 2013). For a full description of the scoring system please see (Appendix 8) In accordance with these systems, possible scores ranged from 0-2 for questions 1 and 2 and 0-1 for question 3 (see section 2.1.4). Using the example in section 2.1.4 of a clip depicting the use of a ‘white lie’ the scoring system was as follows.

For question 1:

2 points: Answers had to identify the main theme of the story (e.g. stating it was a white lie), or an elaborated description that articulated this point (e.g. Max did not want to hurt Alice’s feelings).

1 point: Answers that were partially correct or in which an aspect of the correct description was missing. Examples of these kinds of responses often included solely factual information or simple internal state responses (e.g. he is nice).

0 points: Answers that were irrelevant to the social exchange or incorrect (e.g. he thought it was good).

Psychological-State Talk (PST) was also scored on a scale of 0-2:

2 point: Answers that included 2nd order (or higher) PST e.g. thoughts about thoughts/feelings (e.g. 'he wanted her to think he liked it'), or complex mental states (e.g. 'he's telling a white lie').

1 point: Answers that used 1st order PST e.g. one person's thoughts, or feelings (e.g. 'he's being kind').

0 points: Responses that involved no mention of psychological states.

The scoring schedule for question 2 was developed by the research team as no other previous research was identified using similar methodology. To improve direct comparison between the two questions and as the 0-2 point system has proved successful in generating reliable scoring in novel research (Castelli, et al., 2002; Devine & Hughes, 2013) and neuropsychology more generally (D. Wechsler, 1997, 1999; D Wechsler, 2008) the same system was adopted, but with a slight alteration in emphasis. Importantly, for these responses, the utterance had to make sense with regard to continuing the conversation so therefore had to make sense in light of the intentions of the last speaker.

Again, using the same scenario as described in section 2.1.4. as an example, the scoring system was as follows:

2 points: Answers had to acknowledge the underlying intention of the utterance (e.g. 'thanks, but how do you think I could improve it?')

1 point: Responses were partially correct and/or minimal (e.g. 'thanks').

0 point: Answers were either irrelevant (e.g. do you want me) or incorrect (e.g. misunderstanding the intention of the white lie such as 'don't take the mickey').

Question 3 was scored as either 0 or 1 and correct answers had to mention the relevant piece of information e.g. a guitar in the example in section 2.1.4.

2.5.2 Control clips

The control clips followed the same 0-2 point system devised in previous research using similar control stimuli (White, et al., 2009). In all three control clips a suggested solution was presented to the character that was logical, but varied in its content e.g. paying a bill over the year because it is cheaper. The scoring system was as follows:

Question 1:

2 points: Answers had to articulate exactly why the suggestion had been recommended (e.g. to save them money or it is cheaper).

1 point: Answers that were partially correct (e.g. because they have the money now).

0 points: Irrelevant or incorrect answers (e.g. it's more expensive).

The mental state scoring system was the same as for the experimental clips (see section above) although reference to psychological states was not the focus of the control clips.

Question 2:

2 points: Answers had to acknowledge the reasoning behind the suggestion (e.g. 'yeah that's the best option').

1 point: Answers were deemed partially correct.

0 points: Incorrect or irrelevant answers were scored as 0 (e.g. we could save more money if we didn't pay for the year).

Question 3 was scored as either 0 or 1 and correct answers had to mention the relevant piece of information

2.6 Psychometrics

2.6.1 Item Analysis

Internal consistency was assessed for each scale of the SSFT as shown in Table 2 below:

Table 2: Internal consistency for both the Experimental and Control film clips.

| | Experimental Films | | | | Control Films | | | |
|------------------|--------------------|------|-------------|--------|---------------|------|-------------|--------|
| | Intention | PST | Interaction | Memory | Intention | PST | Interaction | Memory |
| Cronbach's Alpha | .594 | .454 | .745 | * | .323 | .395 | -.232 | * |

Table2: * No variance in the items resulted in too few items being preset to represent the scale adequately.

The internal consistency values for the experimental clips' Intention and Interaction scales were both adequate and satisfactory, respectively. The PST rates however, were low. This is likely to be related to the intentional design of the clips to have varying levels of mental state attribution necessary to understanding them. The Memory questions showed no variance (all participants passed) on 9 items, undermining the alpha statistic. The Control questions showed poor internal consistency on all scales and the Memory questions again showed no variance on 2 items (all participants passed) meaning an alpha statistic could not be calculated. The low alpha in the control films again reflects the intentional variation in control material (economic decisions, understanding of natural phenomena, using left-over food to prevent waste) and the different components of reasoning needed to pass the clips. Therefore, one would not expect adequate alpha values on these subscales as shown on the experimental clips which appear to be tapping an underlying construct.

2.6.2 Inter-rater Reliability

Inter-rater reliability was established by double rating 20% of the responses where the second independent rater was blind to group (shown in Table 3 below).

Table 3: Intra-class coefficients for each dimension of the SSFt for both the ASD and control group.

| | Experimental Clips | | | | Control Clips | | | |
|----------------|--------------------|-----|-------------|----------|---------------|-----|------------------|--------|
| | Intention | PST | Interaction | Memory | Intention | PST | Interaction n | Memory |
| ASD (n=4) | .93 | .86 | .84 | 1 | .93 | 1 | 1 | *(3/3) |
| Controls (n=4) | .99 | .96 | .95 | *(12/12) | .91 | 1 | *(6/6) | *(3/3) |

Table 3: * (number/number) = no variance observed in performance so coefficient could not be calculated so each rater's scores are displayed.

The control group scored at ceiling for memory questions on experimental and control clips. Participants with ASD also correctly answered all memory questions relating to control clips. There was also no variance for scores on interaction questions for control participants. For this reason an intra-class agreement statistic could not be calculated for this data. Given good agreement between the independent rater blind to group and hypothesis and the first author, the latter's ratings were used for the subsequent analyses.

2.7 Phase 2:

2.7.1 Ethics

Ethical Approval for the second phase of the study was granted by the National Research Ethics Service Committee – London, Westminster (13/LO/0092), the Behavioural and Developmental Psychiatry Clinical Academic Group and the Institute of Psychiatry/South London and Maudsley NHS Foundation Trust Research and Development Office (R&D2013/016). Information sheets were provided to all participants before they took part in the study and written consent was taken before the study took place (copies of both documents and the letter of approval can be found in Appendix 9)

2.7.2 Design

A mixed methods design was used whereby both between and within subject variables were analysed. One group consisted of adults with a diagnosis of either Asperger's syndrome (AS) or High Functioning Autism (HFA), which will be collectively referred to as the ASD group. The control group consisted of an age, gender and verbal IQ matched adults. Using the results of Dziobek, et al. (2006) as an estimate, and setting the alpha level at 0.05, a sample size of 20 in each group was deemed sufficient to reach the recommended statistical power of .80 to avoid type 2 error.

2.7.3 Participants

A total of 40 participants were recruited into the study; 20 with a diagnosis of ASD and 20 healthy controls. The ASD group had all been assessed by the Behavioural Genetics Clinic (BGC), South London and Maudsley NHS Foundation Trust, a national specialist, tertiary level adult ASD diagnostic service. Participants were recruited by clinicians working within the BGC or by the researcher if individuals had consented to be contacted for research purposes during previous contact. The control group was recruited through an opportunity sample and advertisements in the local community detailing the research.

2.7.4 Inclusion Criteria

In order to be included in the study, participants in the ASD group had to meet the following inclusion criteria:

- **A formal diagnosis of either AS or HFA provided by a multi-disciplinary team according to ICD-10 criteria.**
- **A reliable informant had completed an ADI-R**
- **If no informant was available, the ADOS had been completed at assessment**
- **Aged between 18 and 65 years at the time of testing.**
- **Fluent in English**

- **Verbal IQ >70 as measured by the Wechsler Adult Intelligence Scale-III/IV (WAIS-III/WAIS-IV; Wechsler, 1997; Wechsler, 2008) or the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999).**
- **No other neurodevelopmental or organic disorder present e.g. no velocardio facial syndrome, hyperkinetic disorder, or brain injury**
- **None of the following psychiatric diagnoses present at assessment: schizophrenia, eating disorders, personality disorder or substance abuse/dependence.**

Inclusion criteria for the control group were (in addition to the criteria above excluding the ASD diagnosis and ASD structured interviews) :

- **No formal mental health or neurodevelopmental disorder diagnosis.**
- **An AQ score below 32.**

Demographics of both groups can be seen in Table 4

Table 4: Demographics for both groups: Mean (SD)

| | ASD N=20 | TD N=20 | t | df | p-value | effect size: d | 95% mean Difference CI . |
|----------------------------|---------------------------|--------------------|-------------------|-----------|----------------|-------------------------------|-------------------------------------|
| Age in years | 30.60 (6.52) | 30.65 (6.27) | -.025 | 38 | .980 | 0.007 | -4.05 – 3.80 ^a |
| Gender Ratio (F:M) | 0:20 | 1:19 | 1.03 ^c | 38 | .31 | 0.33 | - |
| Verbal abilities | 105 (17.01) | 111 (11.52) | -1.35 | 38 | .186 | 0.42 | -15.2 – 2.7 ^a |
| Autism Quotient | 34.2 (7.42) ^b | 15.6 (7.21) | 7.926 | 37 | .001 | 2.52 | 13.91 – 23.4 |
| ADI-R Social Domain | 13.53 (7.86) ^b | - | - | - | - | - | - |
| ADI-R Communication Domain | 11.11 (5.42) | - | - | - | - | - | - |

Table 4 : ^aindependent bootstrap test derived confidence intervals ^bASD N=19 ^cChi squared value.

The two groups were matched for age, gender and verbal IQ and these variables were not significantly different from one another (see Table 4). The AQ acted as a screening measure for ASD traits and the groups were significantly different in their responses, with the ASD group scoring higher (see Table 4). Of particular importance, no individual in the control group scored above 32 on the AQ; the cut-off for significant ASD traits (Baron-Cohen, Wheelwright, Skinner, et al., 2001). In all but one case, a suitable informant was available to provide developmental history information for the participant's diagnosis via an ADI-R. For the individual who did not have ADI-R data, diagnosis was supported by an ADOS. One participant was unable to complete the AQ due to testing constraints.

2.8 Measures

2.8.1 Wechsler Intelligence Tests

The Wechsler Abbreviated Scale of Intelligence (WASI) is a brief reliable and valid measure of general intelligence that is recommended for research purposes (Wechsler 1999). The test consists of two verbal and two non-verbal subtests. Participants complete the verbal subscales (Vocabulary and Similarities), which are required to calculate a verbal IQ score.

In cases where a neuropsychological assessment had been completed within the South London and Maudsley NHS Foundation Trust Adult Autism clinics, participants' verbal IQ scores as defined by the Wechsler Adult Intelligence Scale –III (WAIS-III) short form (Axelrod, Ryan, & Ward, 2001) The WASI and the WAIS-III scores show good convergent validity (Wechsler, 1999). In two cases, the verbal comprehension index of the Wechsler Adult Intelligence Scale –IV (WAIS-IV) was used (Wechsler, 2008).

2.8.2 The Autism Quotient (AQ)

Please see section 2.3.1 for a description of the measure.

2.8.3 The Twenty item Toronto Alexithymia Scale (TAS-20).

The TAS-20 is a 20 item self-report instrument developed to identify alexithymia traits in both clinical and non-clinical populations (Bagby, et al., 1994). The questionnaire uses a Likert scale: 'strongly disagree, moderately disagree, neither agree or disagree, moderately agree or strongly disagree'. Total scores across all items indicate the degree of alexithymia present (see Appendix 10)

The TAS-20 (Bagby, et al., 1994), shows good internal consistency (0.81), test-retest reliability (0.77) and the data were best described using a three factor solution that is theoretically consistent with the alexithymia construct; '1) Difficulties in identifying feelings 2) Difficulties in describing feelings 3) Externally orientated thinking' (Bagby, et al., 1994). The measure has been validated internationally (Taylor, Bagby, & Parker, 2003). Importantly, for the current study the TAS-20 has shown to be a reliable measure of alexithymia in adults with high functioning ASDs and shows good test-retest reliability, convergent validity and discriminate validity in this population (Berthoz & Hill, 2005)

2.8.4 The Interpersonal Reactivity Index (IRI)

The IRI is a 28 item self-report questionnaire designed to test empathy as a multi-dimensional construct (Davis, 1980; 1983; please see Appendix 11 for a copy). Items are equally divided across four factors: 'Perspective taking, Empathic concern, Personal distress in difficult interpersonal situations, and Fantasising about imagined situations'(Davis, 1983). Participants answer each question on a 0-4 Likert scale from 'does not describe me very well' to 'describes me very well'. Scores are totalled for each factor.

In the original paper describing the IRI's development (Davis, 1980) all four factors showed sufficient levels of internal consistency (Cronbach's alpha 0.70-0.78) and test-

retest reliability (0.61-0.81) in both males and females and the factors were stable across two independent samples (total N=1169). It shows appropriate convergent validity and theoretically consistent between-factor correlations i.e. positive correlations between perspective taking and empathic concern (Davis, 1980). The measure has been validated cross-culturally and the psychometric properties are acceptable in the translated versions (Fernández, Dufey, & Kramp, 2011; Siu & Shek, 2005). Pertinent to this study, the IRI has been used in a high functioning adult ASD sample and each factor effectively discriminated the ASD group from a matched typically developing adult sample (Rogers, et al., 2007)

2.8.5 Reading the Mind in the Eyes task (Eyes)

The Eyes task is a forced choice measure of mentalizing abilities that consists of 36 items (Baron-Cohen, Wheelwright, Hill, et al., 2001b). For each item, participants view a photo of a person's eye region. Participants have to then choose (from four options), which word presented best describes the person's expression. Correct scores are then totalled giving a score in the range of 0-36 (for copy see http://www.autismresearchcentre.com/arc_tests.)

In the original study, the Eyes task discriminated between a group of adults with high functioning ASDs (N=15) and three control groups; a community sample (N=122), student sample (N=103) and IQ matched controls (N=14) and performance on the measure was negatively correlated with ASD traits as measured by the AQ in all groups (Baron-Cohen, Wheelwright, Hill, et al., 2001b). The Eyes test has been translated into 24 languages, but the psychometric properties of the measure have rarely been assessed. Internal consistency is mixed, with Cronbach's alpha values ranging from 0.58-0.70 in adult non-clinical populations (Vellante et al., 2012). Test-retest reliability of the measure is acceptable in adult populations (Vellante, et al., 2012; Yildirim et al., 2011) although scores should be considered an 'approximation' due to variance in performance over time (Hallerbäck, Lugnegård, Hjärthag, & Gillberg, 2009). The convergent validity of the Eyes test with self-report measures of empathy is mixed (Muller et al., 2010; Vellante, et al., 2012). A recent meta-analysis suggests the Eyes is a reliable measure to discriminate high functioning adults with ASDs and neurotypical adults (Chung, et al., 2013) and is recommended by experts in the field (Pinkham et al., 2013).

2.8.6 The Awareness of Social Inference Test (TASIT)

Participants completed the forced choice 'Emotion Recognition' subsection of the TASIT (McDonald, Flanagan, & Rollins, 2002). Participants view 28 short film clips, where an actor performed one of the 6 universal emotions: Anger, Sadness, Happiness, Anxiety, Surprise, Disgust, or was emotionally 'Neutral.' Following each clip, participants state the most 'dominant, persistent or pervasive emotion' (McDonald, et al., 2002). Total scores are calculated ranging from 0-28 respectively.

The TASIT in its entirety shows good test-retest reliability ($r = .74-.88$) and good correlation coefficients ($r = .68$) with static second order ToM measures evidencing convergent validity (McDonald et al., 2006). While the TASIT was originally developed for identifying social deficits following traumatic brain injury (McDonald & Flanagan, 2004) its construct validity is supported by it being a suitable measure for identifying social difficulties in patients with schizophrenia (Sparks, et al., 2010) and pertinent to this study and most recently, adults with high functioning ASD (Mathersul, et al., 2013b). Part III is also recommended by experts in the field of social cognition (Pinkham, et al., 2013).

2.8.7 The Frith-Happé Triangles (F-HT)

The F-HT is a non verbal dynamic ToM task (Castelli, et al., 2002). Participants view 1 practice animation and then 4 experimental animations on a computer screen used in Castelli et al., (2002) study. The animations feature a small blue and a large red triangle that interact throughout the clips within a white space and in three of the experimental clips a blue enclosure was present in the centre of the screen. The movements of the triangles depict a range of mental states; 'Surprising, Coaxing, Mocking and Seducing', and the animations were designed such that the triangle's thoughts and feelings could be inferred from their behaviour. The clips last between 36 and 44 seconds. Following each clip the participants are asked 'What was happening in that animation?'. Participants' responses are then written verbatim and coded on a previously established 'appropriateness' scale of 0-2 (Castelli, et al., 2002) and PST scale of 0-2. Scores therefore range from 0-8 for each scale.

The 'F-HT' task has been shown to differentiate between a high-functioning adult ASD group (N=10) and a verbal ability matched adult control group (N=10). The 'F-HT' has been used to assess mentalizing abilities in young children (Abell, et al., 2000) adolescents (Moriguchi, Ohnishi, Mori, Matsuda, & Komaki, 2007), in non-English speaking cultures (Moriguchi, et al., 2006) and has been validated as both an explicit and implicit measure of ToM (Klein, Zwickel, Prinz, & Frith, 2009).

2.8.8 The Strange Stories (SS)

The SS is a written verbal vignette task that has been used extensively in the ASD ToM literature (Dziobek, et al., 2006; Happé, 1994; Jolliffe & Baron-Cohen, 1999; Roeyers, et al., 2001; Spek, Scholte, & Berckelaer-Onnes, 2010). Participants completed a short form of the SS task (White, et al., 2009), consisting of 8 short vignettes (two versions of the following themes: White lie, persuasion, double bluff and misunderstanding). After being read the short vignettes participants explain the character's behaviour or language by answering a single question, usually in the form of 'Why did X [the character] say that?' The vignettes are left in front of the participants to reduce memory loading. Responses are recorded verbatim and are coded on a previously devised 0-2 accuracy scale (White, et al., 2009) alongside a PST scale of 0-2. Total scores therefore range from 0-16 for each scale. For an example of the material and related scoring (see Appendix 12).

In its original form, the SS (Happé, 1994) consists of 24 mental state stories where the actions and/or language of the characters could be best explained in terms of mental states e.g. 'he was *joking*' or 'he did not *want to hurt her feelings*'. 6 physical state stories act as control stimuli where the character's behaviour could be best described in terms of physical actions e.g. 'it is cheaper to buy the multi-pack of drinks'. Scoring was defined using two orthogonal axes; correct vs. incorrect and physical vs. mental state explanations. In the seminal work, (Happé, 1994) found that the ASD group (N=18) gave more incorrect mental state responses when compared to a healthy adult, healthy child and controls with intellectual disabilities. Similarly, the SS showed good convergent validity with first and second order ToM tasks (Happé, 1994). Replications of short form versions of the SS in high functioning adults with ASD have yielded mixed findings with some studies finding differences in accurate mental state attribution abilities (Dziobek, et al., 2006; Jolliffe &

Baron-Cohen, 1999; Spek, et al., 2010) with others revealing no differences (Roeyers, et al., 2001; S. White, et al., 2009).

2.8.9 The Strange Stories Film Task(SSFt)

The same testing settings and instructions were provided to the participants as described in section 2.3.2. Participants view 3 practice clips, two of which were mental state clips and one a control clip. Participants view 15 clips; 12 mental state clips and three control clips, presented in a quasi-randomised order (A). Half the participants viewed order A and the other half viewed the same clips but reversed (order B). Clips last no longer than 27 seconds in length ($M = 17.5$, $SD = 5.83$) and the measures total running time is six minutes and 21 seconds in total. A black screen is present for 1 second followed by the clips number on a black screen for 2 seconds. Participants are asked three questions following each clip (including the three practice clips) as described in section 2.3.2 (for a copy of the test sheet See Appendix 13)

Participants' responses are recorded verbatim. Intention scores range from 0-2 points for accuracy and 0-2 for mental-state talk, totals range from 0-24 for each scale for the experimental clips. For the control clips scores range from 0-6.

Each response to the Interaction question was scored from 0-2 with totals scores ranging from 0-24 for the experimental clips and from 0-6 for the control clips.

Memory questions are scored as either 0 or 1 giving participants a total score of 0-12 for the experimental clips and 0-3 for the control clips.

2.9 Procedure

Testing took place for all participants within a quiet room within the Institute of Psychiatry. Before attending for an assessment participants completed the AQ, TAS-20 and IRI. During the assessment appointment they completed the SS, Eyes, F-HT, TASIT and the SSFt. The order was counter balanced between the SS and SSFt tests, with participants both doing the SS first and the SSFt last or vice versa, which was stratified by order A or B of the SSFt task. The 'F-HT', SS and SSFt tasks were audio recorded. Participants were thanked and reimbursed for their time and those who wished to hear about the findings of the study were put on a list to be contacted regarding the results. If a neuropsychological assessment had not been previously completed then the WASI verbal subscales were completed to provide an estimate of verbal IQ.

2.10 Statistical Analysis

The Kirmilov-Smirnoff test was used as a measure of normal distribution and homogeneity of variance was measured using Levine's test. No significant outliers were identified including the one female participant's scores, so her results were included in the analysis. T-tests were performed to compare between group mean differences on all measures. Sensitivity analysis was performed using an independent bootstrap analysis to test whether the results were robust against deviations from parametric assumptions (Chong & Choo, 2011). The independent bootstrap test is nonparametric. Thus 95% mean difference confidence intervals obtained from the bootstrap test were reported alongside such cases to support outcomes of the test statistic.

ANCOVA was completed, with VIQ as a covariate, in all cases where VIQ correlated with performance on behavioural measures of social cognition. Homogeneity of regression slopes was assessed in all cases supporting the utility of ANCOVA to partial out variance attributed to VIQ. Bootstrap pair wise comparisons were performed in cases where normal distribution was violated. All bootstrap tests were based on 5000 samples.

Alpha values were set at $<.05$ and effect sizes calculated using Cohen's d (Cohen, 1992). Partial Cohen's d effect sizes were calculated for the ANCOVA analysis (Wildt & Ahtola, 1978).

Depending on the variables' parametric status correlations were calculated using either Spearman's or Pearson's correlation coefficient. Furthermore, on the advice of a statistician the alpha value was reduced to $<.01$ to account for multiple comparisons. Trends therefore represent associations $<.05$. In instances where VIQ correlated with the performance on the measure of social cognition partial correlations were completed to control for any variance attributable to VIQ. *R squared* was calculated for coefficients that were significant/trends. To compare whether correlation coefficients were statistically different from one another, Fischer *r*-to-*z* transformation was completed.

A Receiver Operator Characteristic (ROC) curve was performed to demonstrate the traditional social cognition measures and the SSFt's ability to assign participants to their correct diagnostic group. The ROC curve is plotted along two axes: Sensitivity along the Y-axis, which demonstrates the test's ability to correctly identify the ASD participants as belonging to the ASD group (true positives) vs. 1-specificity along the X-axis, which is defined as the degree in which, control participants are incorrectly identified as being part of the ASD group (false positive; Park, Goo, & Jo, 2004).

Cut-offs for each measure are generated that yield differing levels of sensitivity and 1-specificity. The cut-off scores' relative sensitivity and 1-specificity values known as 'operating points' (Park, et al., 2004) are graphically displayed along the sensitivity vs. 1-specificity axis' to form the ROC curve. From this, an area under the curve (AUC) value can be calculated, which is the 'average value of sensitivity for all possible values of specificity' (Park, et al., 2004). The higher the AUC, the more effective the measure is at assigning participants to their diagnostic group. A perfect diagnostic measure would therefore have an AUC of 1, which represents 100% sensitivity (e.g. a cut-off on a particular measure where all ASD participants score below) and 0% 1-specificity (e.g. none of the control participants' score below the assigned cut-off). As these two dimensions are in tension with one another, optimal cut-offs for each measure in this study (see Park, et al. (2004) for scenarios where this may not be the case) therefore represent the largest difference between sensitivity and 1-specificity values.

3 Results

The results will first focus on between group differences on the traditional measures of social cognition, empathy and alexithymia. Next, between groups differences on the SSFT will be assessed. Further investigation into the individual clips from the SSFT will follow and how performances on the SSFT scales are related within each group. Analysis will then turn to the SSFT's ability to discriminate the two groups and the SSFT's convergent validity with traditional social cognition tasks, questionnaire measures of empathy, alexithymia and ASD traits/symptoms.

In the following sections the *F*-value (*F*), *t*-statistic (*t*), degrees of freedom (*df*), Cohen's *d* effect size estimate (*d*) and 95% (mean difference) confidence intervals (*CI*) are displayed with the corresponding abbreviations. *F* always represents an ANCOVA with VIQ as a covariate. Age was not found to correlate with any of the behavioural measures and so was not included in analyses. Where correlation analysis is significant/shows a trend *r* squared will also be reported (*r*²).

3.1 Traditional Social Cognition Measures

Analysis was first conducted to assess whether the two groups performed differently on the traditional social cognition measures, which is shown in Table 5 below.

Table 5: Results by group for tests of social cognition: Mean (SD)

| | ASD N=20 | Controls N=20 | t_{df} | F_{df} | p - value | d | 95% CI |
|--|--------------|--------------------------|--------------------|--------------------|----------------|------|---------------------------|
| Strange Stories | | | | | | | |
| Accuracy ($max=16$) | 11.65 (2.30) | 13.35 (1.73) | | 5.14 ₃₇ | .029 | 0.72 | 0.13-2.78 ^a |
| Psychological State Talk ($max=16$) | 11.65 (1.87) | 12.40 (1.35) | | 1.13 ₃₇ | .294 | 0.34 | -0.49 -1.52 ^a |
| Eyes | | | | | | | |
| Accuracy (max = 36) | 27.68 (4.27) | 25.00(4.08) ^b | | 2.15 ₃₆ | .152 | 0.46 | -0.72 - 4.48 ^a |
| Frith-Happé Triangles | | | | | | | |
| Accuracy ($max=8$) | 3.35 (2.11) | 5.10 (2.17) | 2.14 ₃₈ | | .040 | 0.68 | 0.08 - 2.82 |
| Psychological State Talk ($max=8$) | 2.45 (1.50) | 4.45 (2.31) | 3.25 ₃₈ | | .001 | 1.03 | 0.75 - 3.25 |
| TASIT | | | | | | | |
| Accuracy (max = 28) | 23.75 (2.36) | 24.75 (1.92) | 1.47 ₃₈ | | .15 | 0.47 | -0.29 -2.35 ^a |

Table 5. ^a bootstrap derived confidence intervals. ^b N=19.

The analysis revealed a significant (medium effect size) group difference between the adults with autism and the controls on the SS accuracy scores, but not on the degree of PST used to explain behaviour (see Table 5). While individuals with ASD were using comparable levels of mental state language to controls, their responses were less accurate. Bootstrap analysis was performed, revealing that the findings were robust against violations of normal distribution (see Table 5).

Group differences between the F-HT revealed that compared to controls, individuals with autism were significantly less able to accurately infer the plot of the short animations and the triangles' 'intentions' (see Table 5). The group difference represented a medium effect size (see Table5). Furthermore, they were less likely to imbue the geometric shapes

with thoughts and feelings and use mental state language to describe the interactions of the animated objects. This group difference yielded a large effect size (see Table 5).

Analysis revealed that both groups performed equally on the Eyes task meaning their abilities at inferring mental states from the eye region did not differ. Both groups performed equally on the emotion recognition subtest of the TASIT suggesting that understanding basic emotions presented in video format was not significantly different between groups (see Table 5) Both of these effect sizes were small and confirmed by an independent bootstrap test (see Table 5).

Correlation analysis revealed no significant association between the four standardised measures of social cognition (SS, Eyes task, TASIT and F-HT; see Appendix 14). Additional analysis was completed to compare the association of the traditional social cognition measures to the empathy and alexithymia questionnaires. Only the SS accuracy scores showed a trend towards a significant correlation with the fantasising subscale of the IRI in the ASD group ($r(16) = -.488, p = .04$). No other significant associations were found between the social cognition measures and the questionnaire measures (TAS-20, IRI and AQ) for both groups, and the Social and Communication domains of the ADI-R in the ASD group (See Appendix 15)

3.2 Questionnaire Measures

Participants' scores on the questionnaire measures used to assess empathy and alexithymia are shown in Table 6 below:

Table: 6: Self-reported empathy and alexithymia: Mean (SD)

| | ASD N=19 | Controls N=20 | <i>t</i> | <i>df</i> | <i>p-value</i> | <i>d</i> | 95% <i>CI</i> |
|--|---------------|------------------|----------|-----------|----------------|----------|-------------------------|
| IRI subscales [max=28] | | | | | | | |
| Perspective Taking | 13.21 (6.07) | 17.75 (4.79) | -2.60 | 37 | .01 | 0.83 | -8.08 – 1.00 |
| Fantatising | 13.05 (5.58) | 15.90 (4.72) | -1.72 | 37 | .09 | 0.54 | -6.20 – .50 |
| Empathic Concern | 17.42 (4.10) | 17.75 (2.59) | -0.30 | 37 | .77 | .01 | -2.46-1.80 ^a |
| Personal Distress | 14.42 (5.71) | 10.25 (4.28) | 2.59 | 37 | .01 | .80 | .90 – 7.43 |
| TAS-20: | | | | | | | |
| Identify Emotions [max=35] | 20.58 (5.98) | 15.60 (6.02) | 2.59 | 37 | .01 | .83 | 1.08 – 8.88 |
| Describe Emotions [max=25] | 17.95 (3.46) | 12.95 (5.27) | 3.52 | 33 | .001 | 1.11 | 2.11 – 7.89 |
| Externally Orientated Thinking [max=40] | 23.05 (4.48) | 18.05 (4.44) | 3.50 | 37 | .001 | 1.12 | 2.10-7.90 |
| Total [max=100] | 61.58 (10.07) | 46.60 (11.10) | 4.41 | 37 | .001 | 1.41 | 8.09-21.87 |

Table 6: ^a bootstrap derived confidence intervals. ^b N=19. IRI, Interpersonal Reactivity Index. TAS-20, Toronto Alexithymia Scale.

Significant differences were seen between the two groups on the perspective taking subscale of the IRI. This result suggests a reduced inclination to adopt another's view point in interpersonal situations. There was a trend towards the level of fantasising (i.e. imagining what it is like to be another person) to be lower in individuals with autism compared to control participants, representing a medium effect size (see Table 6). Overall, these results suggest that individuals with autism tended to self-report lower levels of cognitive empathy, and especially in perspective taking, as measured by the IRI when compared to the control group .

The pattern of results was different for the affective empathy subscales. Both the control group and individuals with autism reported equal levels of empathic concern, that is, both groups' 'other orientated' (Davis, 1983) feelings of understanding when observing people in difficult situations was equivalent (see Table 6). However, when answering items pertaining to personal distress, (e.g. the level of distress experienced in difficult

interpersonal contexts) individuals with autism rated themselves as significantly higher (see Table 6), and the observed difference was of large effect size.

The TAS-20 revealed consistent differences between the two groups, representing large effects on each of the subscales and the total scale. Additionally, significantly more of the ASD group (52.6%) reported high alexithymia traits (total score > 60; Bagby, et al., 1994) compared to the control group (20%; $X^2(1,39) = 4.51, p = .03$). This suggests the ASD group showed greater difficulties in identifying and describing their emotions. They also showed a tendency to externally orientate their thinking (e.g. not using emotions when making decisions).

3.3 The Strange Stories Film Task

Analysis was conducted to compare performance of the groups on the SSFt's Intention, PST, Interaction and Memory scales for both the experimental and control clips. No effect of order of presentation was seen across any of the experimental clips. Only the control clips' Interaction question showed an effect of order (See Appendix 16). However, there was not a between group difference and error rates were evenly dispersed across the three clips (see Appendix 16). Results were therefore collapsed across both orders. Results are displayed in Table 7 below.

Table 7: Performance on SSFt by group: Mean (S.D.)

| The Strange Stories Film Task | ASD N=20 | Controls N=20 | t_{df} | F_{df} | p -value | d | 95% CI |
|---------------------------------------|--------------|------------------|--------------------|---------------------|------------|------|---------------------------|
| Experimental Clips | | | | | | | |
| Intention ($max=24$) | 15.60 (3.63) | 18.80 (2.33) | | 8.52 ₃₇ | .006 | 0.95 | 0.85 – 4.45 |
| Psychological State talk ($max=24$) | 11.60 (2.44) | 13.70 (2.47) | | 5.23 ₃₇ | .028 | 0.74 | 0.24 – 3.21 |
| Interaction ($max=24$) | 10.75 (3.42) | 17.10 (4.15) | | 25.05 ₃₇ | .001 | 1.62 | 3.38 – 7.96 |
| Memory ($max=12$) | 11.55 (0.69) | 11.85 (0.37) | 1.73 ₂₉ | | .95 | 0.55 | -0.03 – 0.65 ^a |
| Control Clips | | | | | | | |
| Intention ($max=6$) | 4.10 (1.41) | 4.40 (0.99) | | 0.70 ₃₇ | .79 | 0.08 | -0.59 – 0.81 ^a |
| Psychological State talk ($max=6$) | 0.95 (0.89) | 1.25 (1.16) | 0.92 ₃₈ | | .37 | 0.29 | -0.94 – 0.32 ^a |
| Interaction ($max=6$) | 5.50 (0.89) | 5.60 (0.75) | 0.38 ₃₈ | | .70 | 0.12 | -0.63 – 3.98 ^a |
| Memory ($max=3$) | 2.95 (0.22) | 3.00 (0.00) | 1.00 ₃₈ | | .33 | 0.32 | -.18 – -0.04 ^a |

Table 7 ^a Independent bootstrap derived confidence intervals.

The experimental clips revealed the following pattern of results. Participants with ASD scored significantly lower than controls on the Intention dimension of SSFt (understanding the intentions of the speaker) and the observed effect was large (see Table7).

Individuals with ASD were less likely than controls to use PST (e.g. refer to thoughts and feelings) in their responses to explain the character's behaviour and this difference was of a medium effect size (see Table 7).

Participants' responses to the Interaction question (i.e. what they would say next in that situation) showed a similar pattern of results. The individuals with ASD performed significantly worse than controls (see Table 7) and the effect size of this group difference was large.

The control films yielded a different pattern of results. Both groups performed equally well on the Intention, PST and Interaction questions and these findings were supported by independent bootstrap tests (see Table 7). This suggests that they were equally able to reason accurately about why the speaker said what they said and provide a suitable response to the conversational exchange when not required to 'mentalize'.

No significant group differences were seen on the memory question for experimental or control clips, suggesting that both groups were equally able to recall pertinent information regarding the conversational exchange. Ceiling effects were observed however.

While not an a priori hypothesis, exploratory analysis was performed to identify which themes were contributing to the group differences observed on the SSFt. Descriptions of the profile of groups' performance across different experimental clips' themes are presented. Firstly, the profiles of Intention, PST and Interaction scores were investigated within each group (see Figures 1 and 2). Following this, difference scores for each theme (Control group minus ASD group) were calculated to investigate which themes best differentiated the groups by question type (see Figure 3). For all Figures 1-3 the themes are presented in order of descending Intention scores for the ASD group.

Figure 1: Sum of scores on the SSFt by theme in Control group. Presented in descending order of Intention score for the ASD group

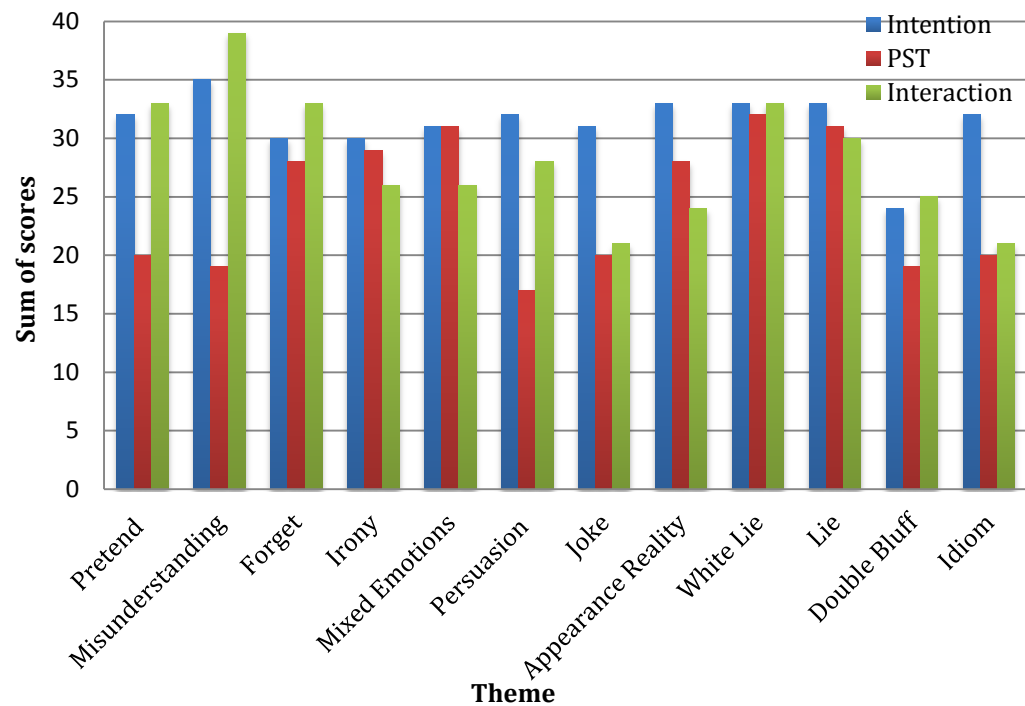


Figure 2: Sum of scores on the SSFt by theme for the ASD group presented in descending order of Intention score.

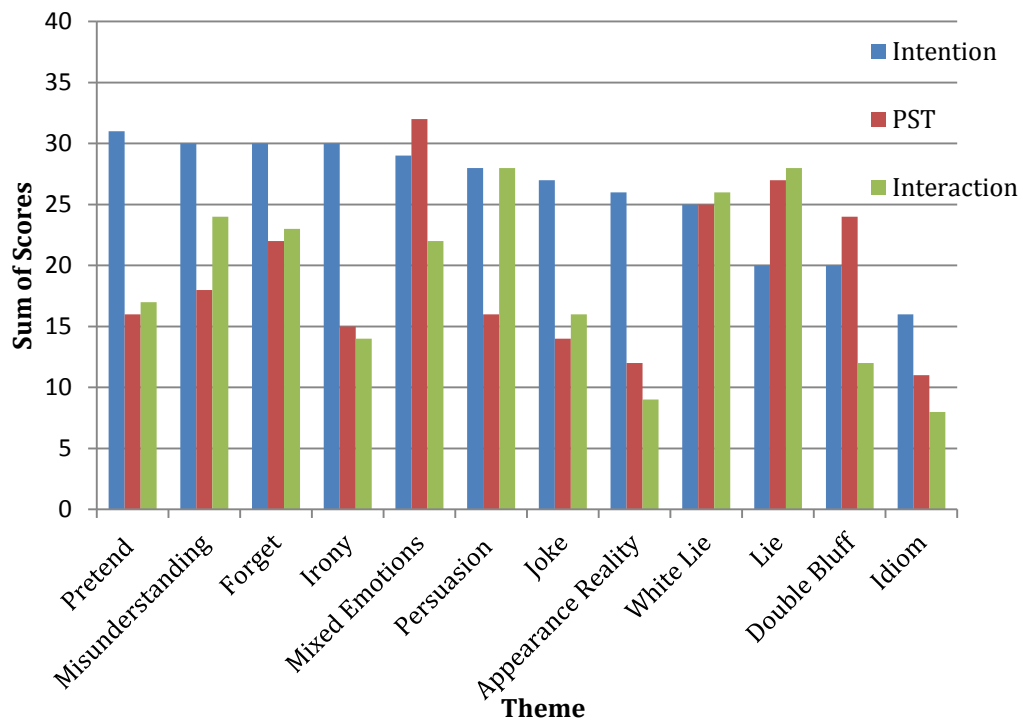


Figure 1 shows the control groups' total scores per scale on each clips' theme. Themes are presented in order of descending Intention scores for the ASD group. All the themes apart from the Double Bluff clip show total scores at or above 30 (average score of 1.5 per individual). The PST scales are at 20 (average score of 1 per individual) or above on all but three clips (Misunderstanding, Persuasion and Double Bluff). The Interaction scores fall above 20 on all of the clips and above 30 on five (Misunderstanding, White Lie, Pretend, and Forget). No clips' scores were at ceiling for any dimension (40).

Figure 2 shows the ASD group's total scores per scale on each clips' theme in descending order on the Intention scale. The Pretend, Misunderstanding, Forget and Irony scenes all yield scores at 30. In contrast the Lie, Double-Bluff and Idiom all fall at either 20 or below. Over half of the clips (n=7) fall below 20 on the PST scale. The reduced rates of PST can be seen in a number of clips where accuracy was high (for example Irony and Joke; see Figure 2). The Interaction scores were generally lower across all items.

Figure 3: Difference scores between control and ASD group on the SSFt by theme in descending order of Intention scores for ASD group.

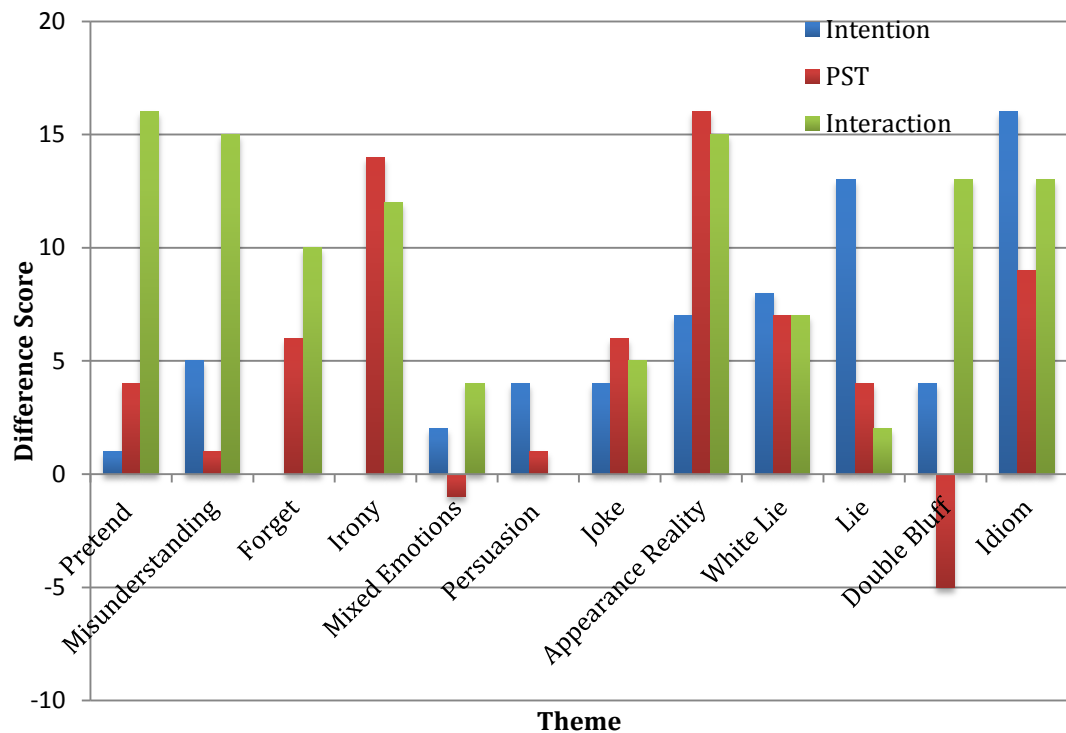
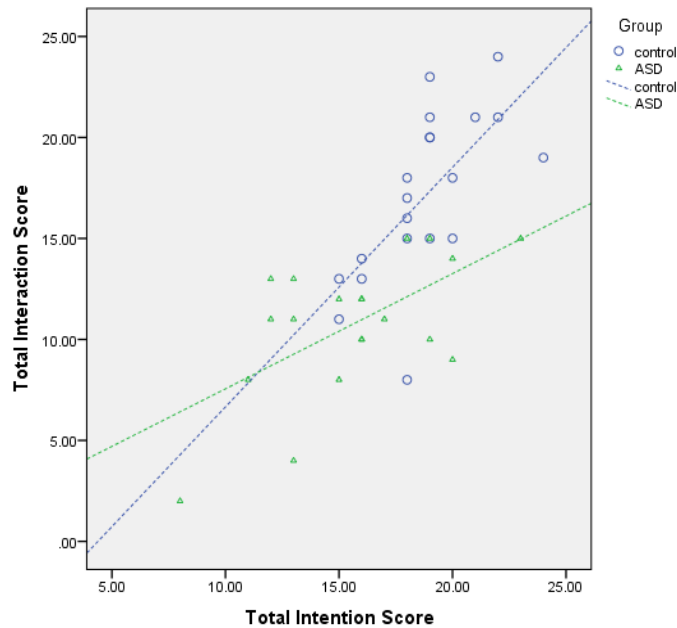


Figure 3 shows the difference in performance by item on SSFt across the three domains (Intention, PST and Interaction) by theme (presented in order of decreasing Intention scores for the ASD group's performance). Positive scores represent better performance by control than ASD group, and negative scores the opposite. The exploratory pattern of results between the dimensions of the SSFt are shown, however since this was not an a priori aim of the study (to compare theme types) no statistical analysis was performed.

The next part of the analysis focused on the relationship between the Intention and the Interaction questions of the SSFt. The relationship between individuals' performance on these two questions is plotted below in Figure 4 with a line of best fit for each group.

Figure 4: Relationship of performance on the Intention and Interaction questions of the SSFt for both groups

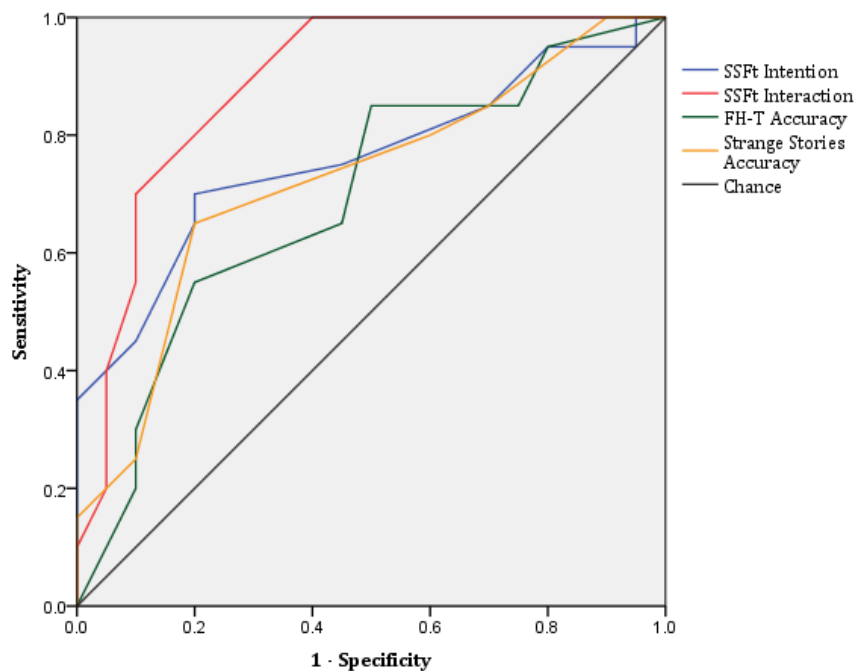


Analysis revealed a trend towards a significant association between the Intention and Interaction scores of the SSFt in the ASD group once verbal abilities had been controlled for ($r(17) = .53, p = .02$). For the controls however this association was statistically significant ($r(17) = .60, p = .007$). Fischer r -to- z transformation revealed that these two coefficients were not statistically significant however ($z = -0.3, p = .76$).

3.4 Group Classification Using the SSFt and Traditional Social Cognition tasks.

One of the main aims of the study was to assess whether the SSFt could discriminate between participants with and without a diagnosis of ASD compared to the other social cognition measures. Figure 5 below shows a ROC curve to assess the sensitivity and 1-specificity of each measure

Figure 5: ROC curves for each social cognition measure and chance.



The ROC curve in figure 5 demonstrates each social cognition measures' ability to accurately assign the participants to their respective diagnostic group. Only measures in which there was a significant difference in means between the two groups were included. PST scores were not included as this scale did not differentiate accurate from inaccurate mental state attribution. The AUC values and corresponding 95% confidence intervals for the scales were .89 (.78 - .99) for the SSFt Interaction question, .76 (.61 - .91) for the SSFt Intention question, .72 (.56 - .88) for the SS Accuracy score and .69 (.53 - .86) for the F-HT. All of the AUC values were significantly above chance ($p < .05$). Overall, this suggests that while all the measures perform above chance, the SSFt Interaction and Intention scales were on average more accurate at correctly assigning participants to their correct diagnostic group. The small sample size in the current study, however, necessitates

caution when interpreting the findings as the confidence intervals for each measure show some overlap.

The preliminary optimal cut-offs and their respective Sensitivity and 1-specificity values are presented below for each measure in Table 8.

Table 8: Cut-off scores for each social cognition measure.

| Test Scale | Cut-off | Sensitivity | 1-Specificity |
|--------------------------|----------------|--------------------|----------------------|
| SSFt Intention | 17.50 | .70 | .20 |
| SSFt Interaction | 13.50 | .80 | .20 |
| Strange Stories Accuracy | 12.50 | .65 | .20 |
| F-H T Accuracy | 3.50 | .55 | .20 |

Scoring below the assigned cut-off would indicate being assigned to the ASD group given the levels of sensitivity and 1-specificity displayed. Table 8 shows how increasing levels of sensitivity, on the SSFt Interaction scale in particular, does not increase 1-specificity value (false positive rate).

3.5 Convergent Validity of the Strange Stories Film Task.

In order to investigate the hypotheses that the SSFt would demonstrate convergent validity with other established measures of social cognition, associations by group were performed (see Table 9).

Table 9: Associations by group Between Performance on the SSFt and Traditional Social Cognition Measures: Spearman's partial correlation coefficient controlling for VIQ (r^2).

| The Strange Stories Film Task | SS Accuracy | SS PST | F-HT Accuracy ^a | F-HT PST ^a | Eyes | TASIT |
|---------------------------------------|-----------------|--------|----------------------------|-----------------------|------|-------|
| ASD [$N = 20$] | | | | | | |
| Intention | .52(27)* | .45 | .04 | -.10 | -.13 | .10 |
| PST | .28 | .37 | -.12 | .06 | .06 | -.38 |
| Interaction | .28 | -.26 | .15 | .07 | -.07 | .35 |
| Controls [$N = 20$] | | | | | | |
| Intention | .19 | .34 | .46(.21)* | .45 | .03 | .32 |
| PST | -.06 | .10 | .35 | .61(.37)** | .08 | .21 |
| Interaction | .17 | .25 | .32 | .36 | .18 | -.01 |

Table 9: Note: Spearman's correlation coefficients except where marked ^a Pearson's * $p < .05$. ** $p < .01$.

For the ASD group a significant association was revealed between the SSFt Intention scores and the SS Accuracy scores, see (Table 9). No other meaningful associations were revealed between these measures (see Table 9). In the control group the SSFt Intention scores showed a trend with the F-HT accuracy and mental state scores. Only the PST variables of the F-H T and the SSFt reached statistical significance and the more stringent level of $p < .01$. No other meaningful associations were revealed between the measures (see Table 9). When comparing the strength of the associations that were significant/trends within group, across groups, only the PST scores on both the SSFt and the F-H T showed a trend to being significantly different between the ASD and control group ($z = 1.89$, $p = 0.06$).

Analysis was then performed to investigate the association between the SSFt and measures of both empathy and alexithymia (see Table 10 below).

Table 10: Association of the SSFt with alexithymia and empathy: Pearson's partial correlation coefficients after controlling for VIQ (r^2).

| The Strange Stories Film Task | TAS-20 | IRI | | | |
|---------------------------------------|--------|------------------|------|-----------------|------|
| | Total | PT | F | EC | PD |
| ASD [$N = 19$] | | | | | |
| Intention | .17 | -.05 | -.34 | .17 | -.01 |
| PST | -.14 | .25 | -.03 | .10 | .41 |
| Interaction | .12 | -.04 | -.07 | .48*(23) | .19 |
| Controls [$N = 20$] | | | | | |
| Intention | -.24 | .50*(.25) | .18 | .42 | .11 |
| PST | -.40 | .56*(.31) | .32 | .32 | -.06 |
| Interaction | -.07 | .23 | -.01 | .43 | .33 |

Table 10 Note: Pearson's correlation coefficients except where marked ^a Spearman's r . * p -value < .05.

The ASD group's performance on the SSFt Interaction question showed a trend to being positively associated with the empathic concern scale. No other meaningful associations were identified (see Table 10). For the control group, a different pattern of results was revealed. The Intention and PST scores of the SSFt showed a positive trend with the perspective taking subscale of the IRI (see Table 10). Of interest, the association was sizeable between the empathic concern scale and the Interaction scale of the SSFt in the controls also, although it was not significant in the current small sample. When comparing the strength of the associations that were significant/trends within group, across groups, only the Intention scores the SSFt and PT scores showed a trend to being significantly different between the ASD and control group ($z = 1.75, p = 0.08$).

3.6 The Strange Stories Film Task's Relationship with ASD Traits/Symptoms

Analysis was then performed to investigate the association of the SSFt and ASD symptomatology (see Table 11 below).

Table 11: Association of SSFt and ASD symptoms: correlation coefficient after controlling for VIQ (r^2).

| The Strange Stories Film Task | ASD Trait/Symptom Measures | | |
|---------------------------------------|---------------------------------|--------------------|--------------|
| | AQ | ADI-R Com | ADI-R Social |
| ASD [$N = 19$] | | | |
| Intention | .01 ^a | -.21 | -.28 |
| PST | -.29 ^a | -.49* (.24) | -.40 |
| Interaction | .12 | -.14 | .02 |
| Controls [$N = 20$] | | | |
| Intention | -.50* (.25)^a | - | - |
| PST | -.60** (.36)^a | - | - |
| Interaction | -.34 ^a | - | - |

Table 11 Note: Spearman correlation coefficients except where marked *Pearson's r . * p -value < .05. ** p -value < .01.

The pattern of results suggests that only the PST scale within SSFt showed a trend towards being inversely associated with ASD symptomatology as defined by the communication domain of the ADI-R (e.g. the higher the ASD symptoms the worse the performance). None of the scales of the SSFt correlated with self-reported ASD traits as measured by the AQ in the ASD group.

In the control group, the Intention scores showed a significant trend towards being negatively associated with AQ scores. The PST scores showed a statistically significant negative correlation with the AQ. No other associations were revealed. When comparing the strength of the associations that were significant/trends within group, across groups, only the SSFt Intention scores and the AQ showed a trend to being significantly different between the ASD and control group ($z = -1.63, p = 0.10$).

3.7 Summary of Main Findings

The overall findings of the study are presented in Table 12 below.

Table 12: Summary of findings across the social cognition measures.

| Measure | Group difference | Association ^a |
|---------|------------------|----------------------------|
| SSFt | ASD<Control | F-H T, SS, IRI, AQ, ADI-R, |
| F-H T | ASD<Control | SSFt |
| SS | ASD<Control | SSFt, IRI (negative). |
| Eyes | n.s | n.s |
| TASIT | n.s | n.s |

Table 12 n.s non-significant, ^a Both trends and significant findings reported

4 Discussion

In the discussion the results will be summarised and then reviewed in light of the study hypotheses. The findings of the SSFt will be discussed in greater depth and contextualised within the current literature. Limitations of the study will be discussed followed by potential future research using the SSFt and its clinical implications.

4.1.1 Group Differences

The primary aim of the current study was to validate a new measure of social cognition designed to discriminate adults with HFASD compared to age, gender and verbal IQ matched controls. The study aimed to overcome a number of the shortcomings in current assessments of social cognition in this group by using dynamic, real-time social exchanges depicting a broad range of linguistic concepts with which individuals with ASD are known to struggle with. The following section will review the results in the light of the study hypotheses:

- 1. The HFASD group will show lower scores on the SSFt's experimental stimuli compared to the control group.**
- 2. The HFASD group and the control group will perform equally on memory questions and all the control stimuli questions of the SSFt.**

These primary hypotheses (H) of the study were confirmed as the individuals in the ASD group showed specific difficulties on SSFt's experimental scales compared to controls, where they were less likely to accurately infer the intentions of the characters, use PST in their responses and provide a suitable response to the social exchange. The pattern of errors observed in the ASD group was unique to scenarios that specifically required inferring an intended meaning from the speaker. Furthermore, these differences could not be attributed to verbal ability. These findings support previous studies that have investigated mentalizing difficulties using dynamic stimuli in HFASD and argued it to be a key feature of the diagnosis (Dziobek, et al., 2006; Golan, et al., 2006; Heavey, et al., 2000). Additionally, the lack of group differences in the films that required logical reasoning to decipher the reason for the speaker's utterance is consistent with vignette based studies of

mentalizing (Fletcher, et al., 1995; Happé, 1994; White, et al., 2009) and suggests that mentalizing is intimately intertwined with non-literal communication.

3. The HFASD group will show lower scores on the SS, Eyes, F-HT but not on the TASIT.

The results partially confirmed the hypothesis that the traditional social cognition measures (excluding the TASIT) would discriminate the two groups. Inferring intentions from vignettes and short animations were therefore challenging which fits with a recent meta-analysis (Chung, et al., 2013). The findings of the current study suggest that basic emotion recognition in dynamic stimuli is not impaired in the HFASD group, which is in keeping with some existing literature (Jones, et al., 2011; Lindner & Rosén, 2006), but inconsistent with other studies (Back, et al., 2007). Variable findings may in part be related to the differing age groups used in the studies (adolescents and children) and different stimuli. Alternatively, basic emotion recognition abilities may be preserved in this high functioning group.

Overall, it appears that basic emotion recognition is intact when all other variables are held constant or when the cue is clear (e.g. shouting indicates the emotion of anger) as in the TASIT. However, when there are contradictory ostensive cues and/or the emotional expression is subtle, HFASD individuals differentially struggle to infer the intentions of the speaker. The current findings did not reveal poorer performance in the ASD group compared with controls on the Eyes task, unlike some previous studies (Chung, et al., 2013), although other studies have similarly failed to find group differences (Roeyers, et al., 2001). Verbal abilities have often not been accounted for when using this measure and recent research suggests the influence of VIQ on this measure is significant (Peterson & Miller, 2012), which fits with the current findings. However, the results of a meta-analysis (Chung, et al., 2013) and expert opinion (Pinkham, et al., 2013) have deemed it one of the leading social cognition measures currently available. Further research is warranted to investigate how much variance is attributable to general cognitive factors and whether the findings in the current study are replicated.

4. The SSFt will be more effective at differentiating between the two groups than the SS, Eyes task, F-HT and the TASIT.

The new measure differentiated the two groups more effectively than previous traditional measures of social cognition. These findings are consistent with studies

showing that more naturalistic measures of mentalizing are more effective at differentiating HFASD samples from controls than traditional ToM measures (Dziobek, et al., 2006; Roeyers, et al., 2001). This finding is particularly relevant to the adoption of the SSFt into future research and clinical assessment with a HFASD population. It suggests that previous tasks may have overestimated the social cognitive abilities of HFASD groups and this may account for mixed findings in the literature investigating mentalizing and social difficulties in ASD (Brunsdon & Happé, 2014). The improved sensitivity and specificity goes some way in justifying the additional administration time of the SSFt compared to the other measures. Finally, the ecological validity of the measure also makes it a more sensitive and informative clinical tool that could be incorporated into neuropsychological assessments, intervention or be used as a social skills teaching tool for individuals with HFASD.

4.1.2 Convergent Validity

- 5. Performance on the SSFt will be associated with the SS, F-H T, TASIT and the Eyes tasks.**
- 6. Performance on the SSFt will be positively associated with the PT and F subscale of the IRI, but not the affective subscales.**
- 7. Performance on the SSFt will be negatively associated with the TAS-20.**

The findings showed that only the F-H T triangles in the control group and the SS in the ASD group (the latter at trend level) were positively associated with the new measure, which partially supported H5. For the controls, visual tasks of ToM were closely associated with performance on the new measure, although visual tasks that had emotion recognition as a focus, whether presented dynamically (TASIT) or statically (Eyes) were not. Therefore, the new task appeared to be related to cognitive rather than emotion processing and this is likely to explain why performance was not related to alexithymia scores (as measured by the TAS-20 and disconfirming H7), but was related to the PT subscale of the IRI (measuring cognitive empathy), partially supporting H6. Alternatively, poor sensitivity of the emotion recognition tasks could explain the lack of association due to minimal variance in task performance.

For the ASD group, traditional verbal tasks of ToM were more strongly associated with the new measure, and again this was independent of emotion processing. This suggests that VIQ and mentalizing abilities are interdependent in able individuals with ASD (Happé,

1995) and that for all the tasks verbal abilities may be in part driving the apparent association rather than mentalizing alone. This will be discussed in more depth in section (4.2.2). It also highlights a weakness in the current measure to tap mentalizing abilities independent of verbal IQ, which has been demonstrated in several previous dynamic social cognition tasks (Dziobek, et al., 2006; Mathersul, et al., 2013b; Roeyers, et al., 2001), but not all (Golan, et al., 2006; Heavey, et al., 2000). However, previous research has not always controlled for the effects of VIQ in associations between new and traditional social cognition measures (Dziobek, et al., 2006). While one has to be tentative in the interpretation of a positive trend, the association between EC on the IRI and the Interaction domain of the SSFt suggests that greater levels of 'affective resonance' may also be associated with this process. Future research could investigate the role of affective resonance and the generation of responses that acknowledge the intentions of others.

- 8. Performance on the SSFt will be correlated with reciprocal social interaction scores and communication scores on the ADI-R in the ASD group, and the AQ in the controls.**
- 9. Performance on the SS, Eyes, F-HT but not the TASIT will be associated with the ADI-R communication and social domains and the AQ.**

Performance on the new measure was associated with ASD symptomatology as measured by the AQ, but only in the control group. The relationship between the Communication domain of the ADI-R and the SSFt showed a sizeable yet non-significant association, partially confirming H8. No relationship was found between the traditional social cognition measures and the ADI-R in the ASD group or the AQ in both groups, disconfirming H9. These findings are inconsistent with the work of Baron-Cohen and colleagues who have shown an association between the AQ and the Eyes task (Baron-Cohen, Wheelwright, Hill, et al., 2001) and other dynamic social cognition measures (Golan, et al., 2006). Previous research has also shown no relationship between traditional social cognition measures and the ADI-R social domain however (Dziobek, et al., 2006), suggesting that informant-reliant childhood symptoms of ASD may not be closely linked to current mentalizing abilities. Furthermore, the childhood literature on ASD symptoms and tasks of ToM is mixed (Bennett et al., 2013) suggesting that cognitive deficits and ASD symptoms may not as be closely linked as some research has suggested (Happé & Ronald, 2008). A limiting factor of the current study is that the small samples per group meant that the findings may represent a type II error as sizeable coefficients did not reach stricter levels of statistical significance and thus had to be considered trends. An important consideration for future research with this group is the inclusion, where possible, of

clinician/informant measures of current ASD symptomatology (e.g. the ADOS) as learned strategies (e.g. verbal) may be masking earlier difficulties and self-report symptoms may not fall in line with behavioural measures. While a self-report measure, the Adult Asperger Assessment (Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005) could be a viable measure of current symptoms alongside the Social Communication Questionnaire (Chandler et al., 2007), which could be included in future research with adults, as the ADOS, while being the gold standard, is a lengthy clinical tool and requires expert training.

10. Performance on the TASIT and the Eyes will be positively associated

11. Performance on the F-HT and the SS will be positively associated.

The traditional measures did not show a significant association with one another or with the cognitive domains of the IRI subscales or the TAS-20 which disconfirmed hypotheses H10 and H11. Previous research has yielded similar findings suggesting that traditional measures of social cognition are not associated with one another (Brent, Rios, Happé, & Charman, 2004; Dziobek, et al., 2006; Lahera, et al., 2014; Roeyers, et al., 2001). This questions the underlying cognitive mechanism that drives individual differences on such measures. The SS, F-H T and the Eyes are all considered ToM measures and the former two have been associated with activation of the 'mentalizing network' (Castelli, et al., 2002; Dodell-Feder, et al., 2011; Fletcher, et al., 1995). The PT subscale of the IRI has been shown to be related to the SS in individuals without ASD (Rogers, et al., 2007). It would seem pertinent to consider why performances on tasks purporting to measure the same underlying construct are not consistently related. One possible consideration is the effects of general cognitive abilities, as while they were not all associated with one another VIQ was associated with performance on the SS and the Eyes. The measures are also different in their demands of the participant, especially between the non-verbal (F-H T, Eyes) and the verbal measures (SS, TASIT). Verbally-based strategies, for example, potentially used to 'hack' through some paradigms may be less fruitful and unable to generalise to other domains (Happé, 1995) resulting in a lack of association. Furthermore, the static (SS,Eyes) vs. dynamic (F-H T, TASIT) difference may result in the tasks recruiting different cognitive processes, and these were driving the lack of association rather than mentalizing per se. The forced choice design (Eyes, TASIT) vs. open ended questions (F-H T, SS) used to assess mentalizing may also be relevant, e.g. forced choice designs provide alternate routes to performance (Johnston, et al., 2008) and are less tasking of generative abilities, which differentially affect those with ASD (Channon, Crawford, Orłowska, Parikh,

& Thoma, 2013). The diversity of the social cognition construct, which is evidenced neurophysiologically (Kennedy & Adolphs, 2012) could also account for this lack of association and the possibility that the different tasks are assessing different components of interrelated, but independent, abilities i.e. 'perceptive-affective vs. cognitive' (Brent, et al., 2004)

4.1.3 Empathy

12. The HFASD group will show lower scores on the cognitive domains of the IRI but not on the affective subscales.

13. Performance on the SS and F-H T will be positively associated with the cognitive domains of the IRI but not the affective subscales.

The IRI revealed that the ASD group reported reduced rates of cognitive empathy when compared to control participants on the PT domain. There were no differences in self-reported affective empathy, partially supporting H12, which is reflected in a range of studies (Dziobek, et al., 2008; Lockwood, et al., 2013; Rogers, et al., 2007). A single trend was revealed between the SS and the F scale of the IRI, although it was negative, in contrast with H13. These findings support the notion that ASD is characterized by distinct empathy profiles that distinguishes it from disorders characterized by difficulties in EC or affective empathy, e.g. psychopathy (Jones, et al., 2010; Smith, 2009a). Knowledge of these differences gives clearer direction for clinicians in developing effective social interventions (Kasari & Patterson, 2012). The elevated levels of PD further support the suggestion that components of affective empathy may be heightened in ASD, which could explain why some social situations are often over-stimulating for such individuals (Smith, 2009b). However, the empathy profile literature is not unanimous as other studies have found reduced affective empathy in individuals with ASD (Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007; Mathersul, et al., 2013b). The use of different scales or combinations of scales as found in Mathersul, et al. (2013b), to measure empathy, may lead to these contradictory findings. Further research is needed to delineate what is driving these cross study differences.

A consistent finding however is that the PT or cognitive empathy subscale is consistently reduced in all studies, which is replicated in the current findings. However, informant- vs. self-report discrepancies found in the adolescent literature (Demurie, et al.,

2011) raise the issue of accuracy in rating empathy and the need for validation research regarding measures such as the IRI to assess PT abilities in HFASD adults. Whether cognitive empathy and mentalizing are synonymous or whether *self-reported* cognitive empathy and *behavioural* measures of mentalizing are related is still left unanswered as the findings in the literature are mixed (Dziobek, et al., 2008; Rogers, et al., 2007). The current negative trend between the F domain of the IRI and the SS suggests that these constructs are not related. Furthermore, the PT subscale of the IRI is deemed the closest to tapping mentalizing (Rogers, et al., 2007) and this was not associated with the traditional measure of ToM (but was associated with the SSFt). Future research is needed to further explore this question as the current study limitations (e.g. small sample) undermine its ability to confirm or disconfirm the cognitive empathy/mentalizing link.

4.1.4 Alexithymia

14. The HFASD group will report higher levels of alexithymia than the controls.

15. Performance on the Eyes and the TASIT will be negatively associated with the TAS-20.

The ASD group did show greater levels of alexithymia, which supports H14 although the lack of meaningful associations between any of the measures and the TAS-20, including those that had explicit emotion processing demands (TASIT, Eyes), does not support the alexithymia hypothesis of the emotion processing difficulties in ASD (Bird & Cook, 2013). However, this may be affected by the small variation in performance on the TASIT making associations difficult to reveal. The findings do not support the claim that alexithymia should be used as a matching variable akin to VIQ (Bird & Cook, 2013) when assessing measures that have a stronger cognitive focus e.g. the SSFt, rather than emotion processing tasks (Lockwood, et al., 2013).

4.2 The Strange Stories Film Task

The SSFt was a novel task developed by the research team for this study. The aim of the project was to create a measure that could detect real world social exchange difficulties in a HFASD sample.

4.2.1 Understanding the Intention question

The group differences on the Intention question of the SSFt can be seen as evidence for a mentalizing impairment in ASD, which has a wealth of empirical support at the behavioural and physiological level (Happé & Frith, 2013). Differences have not always been found between HFASD individuals and controls on ecologically valid measures (Ponnet, et al., 2005) so this measure provides a potential tool to tap subtle social cognitive difficulties within this group. The findings that the control group's performance was not undermined by ceiling effects and the sizeable association between the AQ and the SSFt's Intention question also supports the utility of the measure in identifying individual differences in mentalizing in the general population (see section 4.5).

A potential interpretation of the reduced Intention scores is that impairments in EF may explain the group difference. Holding the necessary information in mind in the short term, updating or inhibiting responses could lead to the differences in the ASD group as such difficulties have been observed in this population (Hill, 2004) especially in more open-ended or naturalistic tasks (Hill & Bird, 2006). Moreover, EF have also been implicated in the developmental trajectory of mentalizing abilities in ASD in childhood (Pellicano, 2012) and into adulthood (Dumontheil, Apperly, & Blakemore, 2010).

While the EF account cannot be falsified per se due to the lack of any well-validated executive function tasks included in the design, the introduction of control clips challenges this interpretation somewhat. The control clips were explicitly designed to require varied knowledge to pass, include equally complex language, be of equal length to the experimental clips and engaged working memory in some instances (e.g. mathematical calculations). To answer correctly participants had to be able to recruit a range of general

cognitive faculties, which makes the isolated group difference observed between the experimental clips puzzling, if one adopts this standpoint. While it is unlikely therefore, that the experimental group difference was solely a result of EF impairments, future research that systematically controls for the role of these processes would be an important avenue of enquiry.

4.2.2 Verbal Abilities and the Strange Stories Film Task

Verbal abilities were associated with the social elements of the SSFt and this finding has been well documented in the social cognition literature in both children (Happé, 1995) and adults in dynamic tasks of social cognition (Golan, et al., 2008). The results certainly suggest that adults are able to infer mental states, but adults with HFA may use verbally mediated compensatory strategies to explicitly decipher the social situation rather than an 'intuitive' reading of it (Bowler, 1992). Correlation does not equal causation, so caution is necessary in extrapolating beyond these findings; however, verbal abilities may provide individuals with HFASD an alternate strategy to generate a verbal response that explicitly states that an intention has been represented (Lind & Bowler, 2009). The implementation of compensatory strategies is a potential hypothesis as to why individuals can pass laboratory based tasks of social understanding, where verbal abilities are [as shown in the current study] associated with performance, yet these high verbal abilities do not translate into real-world adaptive behaviour (Klin et al., 2007).

The demands of the social world are multi-faceted and fast paced. Relevant information, therefore, has to be processed and irrelevant information ignored to accurately infer meaning (Sperber & Wilson, 2002). These explicit verbal strategies may differentially affect those with ASD when reasoning about mental states in such an environment, due to an overall slowing in processing speed when other cognitive demands are high (Chevallier, Noveck, Happé, & Wilson, 2011). A limiting factor of the current research is that compensatory strategies were not directly assessed. For the most part, reaction time is the only proxy measure used to assess the presence of a potential compensatory strategy (this in itself is a limitation and will be discussed in the future research section). If response latencies are greater for individuals with ASD it is hypothesized that the individuals may not be processing the social information as automatically as controls, especially if reaction time differences are only present when

mental state representation is necessary to pass. The research measuring reaction time in social cognition paradigms have yielded mixed findings with individuals with HFASD both in adults (Heavey, et al., 2000) and in children (Chevallier, et al., 2011) and between vignette based (White, Coniston, Rogers, & Frith, 2011) and dynamic tasks using actors or cartoons (Beaumont & Sofronoff, 2008; Heavey, et al., 2000).

While it was not assessed in the current study, the author's impression was that for individuals with ASD, processing speed was reduced particularly for the Interaction question and that testing sessions were longer overall. This hints at the possibility of increased response latencies, although this crude impressionistic assessment is by no means sufficient evidence for the presence or absence of an alternative strategy recruited by the HFASD group for completing the task. The neuropsychological literature points to a cognitive style of ASD being slow but accurate on cognitive tasks (Johnston, Madden, Bramham, & Russell, 2011) so capturing response latencies would be an interesting avenue for future research using this measure. It would be particularly interesting to capture latency by accuracy interactions with the current task. If individuals with ASD were slower (the author's impression) but still inaccurate (the findings), it would suggest that potential strategies were affording no benefit to the individuals when assessing real time social understanding. This would sit in contrast to vignette based measures, which have found greater response latency time in association with increased accuracy (White, et al., 2011), lending greater support for the ecological validity of the task.

4.2.3 Understanding the Interaction Question

One particularly novel element of the current study was asking participants 'if you were in x's situation what would you say next?' or the Interaction question. The wording was considered in depth in order to reduce as far as possible demands on imagination which is known to be affected in ASD (Lord, et al., 1989). Participants had to therefore generate a response that was their own, regarding the situation, taking into account the perspective and information available to the character. The filming, using first person perspective camera work and was intentionally designed to simulate a real life conversation to further facilitate responding to this question. Clearly the design is by no means approaching the naturalistic quality of the empathic accuracy studies where ASD participants have to infer intentions in conversational partners (Ponnet, et al., 2005)

however, the current study overcomes difficulties with vignette based paradigms used in previous research studies that have assessed individuals with ASD traits' responses to 'what would you do next?' (Jameel, et al., 2014).

The Interaction question yielded the greatest group difference for the SSFt, showed a satisfactory degree of internal consistency, and differences could not be attributed to verbal abilities. One explanation is that the findings could be seen as a bi-product of poor performance on the Intention question. It would not be surprising if individuals who struggled to infer the meaning of the speaker then provided answers that did not respond to this communicated meaning. In terms of a ToM account of the social difficulties observed in ASD, one would expect the scores on the Intention questions to predict Interaction question scores (Frith, 1994; Hill & Frith, 2003). Results partially supported this prediction, as descriptive analysis showed some clips had similar Intention and Interaction difference scores between groups (e.g. White Lie, Joke, Mixed Emotions) although this was not universal across the clips. Additionally, there was a significant correlation in the control group and a trend in the ASD group between the Intention and Interaction scores. However, as Intention scores increased in the ASD group they were not associated with higher Interaction scores as seen in the control group, and the Interaction scores for several clips/themes were much lower in the ASD than control group despite similar Intention scores. These findings suggest that difficulties on the Interaction question went beyond secondary problems due to poor attribution of intentions. Mentalizing abilities could therefore be conceived as *necessary but not sufficient* for Interaction abilities in the individuals with ASD.

Again, executive functioning differences between groups may have a potential role, particularly as verbal generation impairments have been observed in adults with HFASD (Hill, 2004). The absence of group differences for the interaction question in the control clips suggests that poor verbal generation alone cannot account for the group differences. A caveat to this argument however, is that the Interaction questions for the controls were often related to a structured problem where a suggestion was made that the participants had to respond to (e.g. paying a bill monthly or every year). *Choosing* appropriate solutions to problems has been shown to cause little difficulty for individuals with HFA compared to *generating* the solution to the problem in a forced choice paradigm (Channon, et al., 2013). Overall, if generation abilities alone were to account for the differences between groups then one would expect the control films to yield group differences and

that was not the case in the current study. That said, ceiling effects on the control clips were observed in both the ASD and control group, meaning true group differences may not have been detected. This is a limitation in most social cognition paradigms (Heavey, et al., 2000) that this study has therefore only partially addressed

Possible insights into the Interaction questions group difference may come from the moral reasoning literature, which has shown a decoupling of intentional understanding from punishing behaviour and moral responsibility (Buon, et al., 2013). In this literature the outcome of the situation seems to bear more weight for individuals with ASD compared to the mental states the situation is contingent upon (Moran, et al., 2011); 'It's *not* the thought that counts'. The Forget scene in the current study highlights this. Individuals with ASD were more likely to see the act of forgetting as the character's fault and/or become angry because of the outcome, thus assigning more responsibility to a mistake. In addition, research has shown that moral behaviour in ASD is guided more by rules than the intentions or internal states of others (Zalla, et al., 2011).

Within the Persuasion scene the responses of individuals with ASD are similar to controls on the Interaction question. In this scene, a pregnant character wants her partner to give up smoking. A clear rule of 'don't smoke around others' or 'don't smoke around pregnant women' could be applied to generate a suitable interaction without directly integrating the partner's '*wanting*'. In this scenario, the behavioural outcome would be the same. Similarly, a 'stock phrase' or 'social script' account may prove a helpful heuristic in understanding the preserved Interaction scores in some circumstances, but not in others (Volden & Johnston, 1999). The recruitment of such scripts has been posited as a possible explanation as to why individuals with ASD appear to do well at inferring the thoughts and feelings of others in generic conversations (Ponnet, et al., 2005), but may struggle when a conversation is more chaotic and lacking the predictable structure of a known topic areas (Ponnet, et al., 2008). Conversely, it could be this overreliance on rule-based responses that makes the generation of the Interaction question responses more challenging for the individuals with ASD in situations that are less clearly defined.

Conceptualising the comprehension of intentions as being necessary, but not sufficient, for social interaction and in turn peer relations is highlighted in recent work commenting on 'social acting' in individuals high in ASD traits (Yang & Baillargeon, 2013). 'Social acting' refers to the use of speech that is intentionally decoupled from the individuals beliefs that

is usually for the benefit of the recipient e.g. white lies (Yang & Baillargeon, 2013). Yang and Baillargeon (2013) propose that it is the lack of 'social acting' and the differential appraisal of its utility that predicts peer relation difficulties seen in adults with ASD traits rather than deficits in mentalizing abilities. It would be interesting in future research to collect qualitative data about the opinions of participants with ASD regarding the characters' behaviour and use of figurative language in conjunction with their evaluation of its utility and how both of these relate to social success or difficulty. From Yang and Baillargeon (2013) work, even if HFASD participants were able to conceptualise why an individual was using figurative language (e.g. to not hurt the other's feelings), their appraisal of its usefulness and subsequent response (e.g. why did you say it's good when you clearly don't believe it?), may undermine the intentions of the individual.

Developmental psychology literature also points to a feedback system between reduced mentalizing abilities and poor peer relations (e.g. children low in ToM are not popular play-mates) (Banerjee, Watling, & Caputi, 2011). This association, however, is mediated by pro-social behaviour (Caputi, Lecce, Pagnin, & Banerjee, 2012), which also predicts peer-relations in adults better than mentalizing abilities (Jameel, et al., 2014). Taken together, future research using the SSFt could investigate the role of current peer relations (or lack thereof) and/or negative past social experiences, which may affect the ability to interpret subtle intentions and provide a suitable response to the social exchange.

The current findings therefore are in support of accurate intention representation being necessary, but not sufficient for appropriate Interactions and a distinct advantage of the current research is the increased sensitivity of the current measure to identify intention inference dis/abilities in a HFASD sample. From this sensitive platform, difficulties in social interaction independent of metalizing may also be investigated.

4.3 Limitations of the study

4.3.1 Generalisability

The study had a number of limitations. The age range of both groups was fairly limited, with the samples being centred on young adults. While age did not correlate with performance on the social cognition measures, extrapolating the findings to an older or younger cohort may be unwarranted. Testing a wider age range would be a next step in the validation process of the new SSFt measure.

Sampling bias is a possible limitation. The recruitment methodology and testing constraints also limited recruitment to individuals who were living either near to or commuting to London. They also had to be motivated to attend the session, willing/able to sacrifice the time, and interested in taking part in research. The information sheets explicitly stated 'social understanding' so participants who responded may have had an interest in this area. These motivational factors may have biased the findings, although it is likely that better adjusted individuals (e.g. without additional social anxiety) took part, so that group differences are perhaps more likely to be diluted than exaggerated.

No data were collected on relationship, living circumstances or occupational status so the groups cannot be assumed to be matched on these variables. Furthermore, ethnicity was not controlled for in the groups and this may be an area of interest given cross-cultural differences in normative social behaviour. In addition, gender ratios were not equivalent in the current study to ASD population estimates of 3.3:1 (males:female) (Simonoff, et al., 2008), as all but one of the ASD participants was male. The female participant was not an outlier on any task, but the lack of any female participants in the ASD group may have underestimated the ASD groups' performance (Baron-Cohen, 2002).

The sample size was small compared to some recent assessments of social cognition with adults with HFASD (Mathersul, et al., 2013b; Spek, et al., 2010) although it was comparable to other studies using the standard measures of mentalizing abilities in adults with HFASD (Dziobek, et al., 2006; Heavey, et al., 2000; Rogers, et al., 2007). Importantly, the group differences found in Dziobek, et al. (2006) research suggested that the current study was sufficiently powered to detect group differences of the anticipated magnitude.

Order effects were also observed on one dimension of the SSFt (Control clips' Interaction question). While the differences were not isolated to the first or last clip and were not different between groups e.g. the ASD group and not the controls were affected by order, the presence of an order effect is likely to reflect participants in each group rather than order. The use of a randomisation procedure to order could overcome this limitation in future research.

Because of the exploratory nature of the study and the design and inclusion of a completely novel task, many variables were examined and there were multiple statistical comparisons. The number of tests in the association analysis necessitated reducing the alpha level to control for type 1 error. The study is likely to have been underpowered to detect differences at this stringent alpha level in many cases. The small sample size is also likely to have increased the confidence intervals in the ROC curve's AUC values for each social cognition measure, making comparisons between the tests sensitivity/1-specificity difficult to assess. Finally, the small sample further undermined the studies ability to test between group differences on the strength of associations across/within measures. These statistical limitations highlight the preliminary nature of the findings and that caution should be exercised when considering the utility of the SSFt. Future research with larger samples would overcome these limitations.

One individual's diagnosis was not confirmed using ADI-R as a reliable informant was not available to provide a developmental history. While his diagnosis was provided by an expert clinician in an outpatient national-specialist tertiary service (within which all the participants were diagnosed) with the support of an ADOS, this lack of informant-based developmental history is a limitation.

4.3.2 The Strange Stories Film Task

The measure itself was limited for a number of reasons. Firstly, the Intention domain did not reach the satisfactory Cronbach's alpha level of 0.7 suggesting that the inter-item correlation was adequate but not satisfactory, which is also highlighted in the low alpha level in the PST variable. While this reflects that the measure may not assess a single underlying construct (Devine & Hughes, 2013), the test was designed to have items with varying levels of difficulty (first and second order ToM) and expected PST complexity, and

this is likely to have added to the somewhat low rates of internal consistency. The control films' low rates of internal consistency across all the domains could be argued as a limitation, although these films were not designed to tap a unitary underlying construct and were few in number (number of variables being related to alpha).

Ceiling performance on the Memory questions on both the experimental and the control clips suggests the SSFt is potentially affected by the same problems as previous measures in not having sufficiently challenging questions to control for general cognitive abilities (Dziobek, et al., 2006). Future work should establish the psychometric properties of this novel test, including test-retest and split half reliability.

The current version of the SSFt and this study was not designed to examine performance differences across distinct types of non-literal utterance. Since, due to time constraints, only one clip was included per theme, theme effects could not be distinguished from clip idiosyncrasies. Future work using multiple clips per theme, normed for difficulty, could test the hypothesis that different themes differ in the degree of mental state inference necessary. For example, understanding an utterance arising from a misunderstanding (e.g. she thought he was a burglar) is arguably simpler in terms of mental state representation than double-bluff (e.g. she *wants* him to *think* she's joking, so tells him the truth because it is so unbelievable). However, as there was only one item per theme, it is impossible to isolate the scores of the theme as being related to the linguistic concept per se as demonstrated in Happé's (1993) early work comparing similes, metaphors and irony to ToM comprehension.

The development of the measure may also be conceptually limited by the methodology of consensus between researchers with regard to what is considered an accurate interpretation or response. Johnston, et al. (2008) provides an insightful yet critical appraisal of the methodology of consensus to determine the most accurate description of a social stimulus. They use the Eyes test as a case in point, where the *true* emotion of the pictured individual is not identified, but decided upon by the authors. While the current study overcomes this limitation in so far as the actors were directed and the emotion agreed upon, the scoring overall may have been biased by the researchers' own cultural opinions of what constitutes an 'irrelevant' response. This criticism applies to all social cognition measures using actors (Devine & Hughes, 2013; Dziobek, et al., 2006; Golan, et al., 2008; Golan, et al., 2006; Heavey, et al., 2000) and agreement between (neurotypical)

raters to score responses, so is not unique to the current research. However, it questions the objectivity of the measure and calls into question the pursuit of objectivity in this line of research. Leading from this, Milton (2012) argues that the ToM hypothesis places the social deficit within the individual, which misrepresents the relational context within which social exchanges occur. He uses the term 'double empathy problem' to highlight that 'the social difficulty' is bi-directional in so much as it resides in both the ASD individual and those without the diagnosis; if researchers with ASD had devised the scoring system would the same 'irrelevant' responses have being coded as such or this kind of research been pursued in the first place? Such theoretical critiques raise interesting considerations, with regard to the nature of and direction of future research in the field of social cognition.

What appears to be relatively uncontentious is that interaction between agents (Zaki, Bolger, & Ochsner, 2008), contextual effects and the use of ever more sophisticated tools that reflect real life interactions are essential for assessing social cognition; this piece of research is a small step in that direction.

4.4 Clinical Implications

The *Autism Act 2009* was a landmark change in UK legislation; restructuring service provision to accommodate the needs of adults with ASD became a legal duty. The accompanying policy documentation that followed the Autism Act 2009 highlights a strategy to better equip society to understand, value and support adults with ASD in the UK (Social, Care Local Government and Care Partnership Directorate, 2014). Health care professionals have a duty of care to use measures that provide an accurate description of individuals with ASD's strengths and difficulties. The current research offers clinicians a possible tool to use in the assessment of mentalizing difficulties in adults with ASD. It is relatively short to administer and has a reliable scoring system. The Interaction question also gives a unique insight into whether the individual's difficulties generating social responses extend beyond their mentalizing dis/abilities. This provides clinicians with a clearer direction for targeted interventions or support for individuals with HFASD. While the existing evidence for the link between training studies of ToM and real world perspective taking abilities is not promising in children (Begeer et al., 2011; Kasari & Patterson, 2012), or adults (Palmen, et al., 2012), the use of sensitive measures is of potential benefit in assessing changes in mentalizing abilities in already high functioning

individuals over time. Such interventions would be particularly beneficial for those individuals that seek relationships yet do not have the skills to maintain them, which lies in contrast to the lay perception that individuals with ASD are not socially motivated (Huws & Jones, 2010). Test re-test reliability would be essential to verify before this next step could be taken.

Mental health problems have been shown to be elevated in ASD populations (Simonoff, et al., 2008) with social anxiety being one the more common conditions (White, Oswald, Ollendick, & Scahill, 2009). Furthermore, awareness of social difficulties in a higher functioning group may be elevated, creating a vicious cycle exacerbating difficulties of anxiety and social impairments (White et al., 2013). Recent studies have shown that treating co-morbid anxiety in HFASD with modified CBT protocols has a significant impact on anxiety symptoms (Lang, Regester, Lauderdale, Ashbaugh, & Haring, 2010;. Storch et al., 2013) and can reduce core social difficulties (Sze & Wood, 2008). The new measure could be used to assess mentalizing abilities before and after therapeutic interventions for social anxiety in adults with HFASD, to assess how interrelated these two constructs are.

4.5 Future Research

The preliminary nature of this study calls for future research to first and foremost support the construct validity and reliability of the measure. The findings are encouraging; however, replication studies will be essential to overcome a number of the statistical limitations of this study, such as small sample size.

Administration and scoring of the SSFt is not overly time consuming, however clinicians' time is often scarce. The participants' responses from the current open-ended questions could, in the future, be used to create a forced choice version of the measure. Alongside the correct responses, distracter answers could include purely 'physical' answers, incorrect answers that include mental state terminology and partially correct answers. Both accuracy rates and the nature of errors between groups could then be analysed (Montag, et al., 2011). Additionally, if the answers were presented electronically, response time could be measured. Furthermore, such a task could be uploaded onto an online platform allowing the possibility of reaching a broader demographic and larger sample. The sensitivity of the current measure suggests that mentalizing difficulties/skills

in the normal population could be assessed and an alternate platform would potentially open up these research opportunities.

There are of course pros and cons to a forced choice design. Research in the problem solving literature has shown intact abilities in ASD individuals abilities to chose appropriate solutions to day-to-day problems with concurrent difficulties in generating them (Channon, et al., 2013). Furthermore, proving participants answers provides an alternative route to completing the task as highlighted by the work of Johnston et al. (2008). In all, a forced choice version would be an exciting next step for the SSFt although careful consideration of the limitations will be essential to the development of an alternate version.

The lack of association between the PT domain in the ASD group and performance on the social cognition behavioural measures, calls into question whether they are tapping the same cognitive empathy construct as being measured by the IRI. It also calls for additional research in the field of insight into mentalizing abilities in adults with HFASD. The literature suggests that self-report measures are valid in HFASD sample for emotion processing (Berthoz & Hill, 2005). With regard to personal reflection on PT difficulties, informants may prove helpful in unpicking whether ASD participants and/or informants are accurately reporting PT abilities (Demurie, et al., 2011) or whether the two constructs are really not as related as previous research has assumed (Rogers, et al., 2007).

The reliance on verbal and behavioural responses makes the measure an explicit task of mentalizing, so a social orientating deficit account cannot be falsified, the ASD group were not attending to relevant social cues for example (Klin, et al., 2003). Future research may be able to explore this hypothesis further with the use of eye-tracking data, which has yielded differences in explicit vs. implicit strategies in HFASD when responding to social cognition tasks (Senju, et al., 2009). It would be interesting to evaluate whether different parts of the face (e.g. eyes vs. mouth) are attended to (Klin, Jones, Schultz, Volkmar, & Cohen, 2002) and how this related to performance in decoding the intended meaning of the speaker and subsequent responses to interaction questions .

While age was not associated with performance on the measure, the demographic range was relatively small. It would be of interest to assess whether the measure could be used in a range of demographics including adolescents and older adults. The

developmental trajectory of mentalizing abilities has been studied most in childhood, with mentalizing capabilities in adolescence being relatively under-researched despite the demands of social cognition being extremely high during this developmental stage (Blakemore, 2008; Happé & Frith, 2013). Research has revealed that mentalizing abilities are continually developing across late adolescence into adulthood (Dumontheil, et al., 2010). Better understanding of the aging process of mentalizing in ASD would be an interesting line of research, and the SSFt might be used to explore its interaction with peer relations, isolation, quality of life and health and wellbeing.

Qualitative methodology may be an avenue of potentially fruitful research. Creation of the Interaction question in the SSFt resulted from a conversation with an associate of the author who was diagnosed with HFASD in adulthood. A theme of this conversation was that the generation of an appropriate response was challenging even when he knew why people were saying what they were saying. Insights from the autobiographies of individuals with HFASD have taught the research communities substantially about the talents and difficulties of HFASD individuals (Frith, 1991). Exploring individuals' subjective experience of completing the measures would be extremely helpful in understanding what strategies they use to understand the situations. Were they using cues from the films and/or past experiences to guide responses? Did they find the scenarios believable and/or could relate to the situations the character's found themselves in (e.g. looking for a romantic partner while being in a relationship)? Did they find the acting or the emotions unbelievable and did this affect their ability to resonate with the characters? These are all unanswered questions that could be addressed with a thematic analysis of participant's understanding of the test itself.

A recent theory has suggested that it is potentially the implicit understanding of what a task is attempting to test and what is expected by the experimenter that may differentiate controls and those with ASD in the domain of EFs (White, 2013). Thus, the mentalizing difficulties observed in the ASD group in the current study may also be impairing the ASD individuals from the onset with regard to what is expected of the task, especially as the questions were left as open-ended and naturalistic as possible.

4.6 Concluding Remarks

The current study developed a novel, dynamic, video-based measure to assess mentalizing abilities in HFASD adults. HFASD adults' performance was characterized by fewer accurate inferences, less psychological state language and a reduced ability to generate a suitable next comment in an interaction where character's intentions were communicated through non-literal language. Intention scores were observed to be necessary, but not sufficient, for good generation of conversational contributions, which most effectively differentiated the two groups. The new task showed convergent validity with traditional social cognition measures, cognitive empathy and ASD symptomatology, but for the ASD group these associations were often a reflection of verbal abilities. Alexithymia was not associated with performance for either group although it was elevated overall in the HFASD sample.

This study suggests that the mentalizing difficulties observed in HFASD are but one piece in a large and complex puzzle of social functioning. However, the current task provides clinicians and researchers with a more sensitive tool to assess social insight, which may be helpful in planning and measuring the effects of interventions.

5 References

- Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, 15(1), 1-16.
- American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Arlington, VA: American Psychiatric Publishing.
- Axelrod, B. N., Ryan, J. J., & Ward, L. C. (2001). Evaluation of seven-subtest short forms of the Wechsler Adult Intelligence Scale-III in a referred sample. *Archives of Clinical Neuropsychology*, 16(1), 1-8.
- Back, E., Ropar, D., & Mitchell, P. (2007). Do the Eyes Have It? Inferring Mental States From Animated Faces in Autism. *Child Development*, 78(2), 397-411.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556-559.
- Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23-32.
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*, 368(9531), 210-215.
- Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and “Reading the Mind in the Eyes”. *Intelligence*, 44(0), 78-92.
- Banerjee, R., Watling, D., & Caputi, M. (2011). Peer Relations and the Understanding of Faux Pas: Longitudinal Evidence for Bidirectional Associations. *Child Development*, 82(6), 1887-1905.

- Barnes, J. L., Lombardo, M. V., Wheelwright, S., & Baron-Cohen, S. (2009). Moral Dilemmas Film Task: a study of spontaneous narratives by individuals with autism spectrum conditions. *Autism Research*, 2(3), 148-156.
- Baron-Cohen, S. (1989a). The Autistic Child's Theory of Mind: a Case of Specific Developmental Delay. *Journal of Child Psychology and Psychiatry*, 30(2), 285-297.
- Baron-Cohen, S. (1989b). Joint-attention deficits in autism: Towards a cognitive analysis. *Development and Psychopathology*, 1(03), 185-189.
- Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd ed.)* (pp. 3-20). New York, NY, US: Oxford University Press.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6(6), 248-254.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, 21(1), 37-46.
- Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, 34(2), 163-175.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241-251.
- Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): A Diagnostic Method. *Journal of Autism and Developmental Disorders*, 35(6), 807-819.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5-17.

- Beaumont, R., & Sofronoff, K. (2008). A New Computerised Advanced Theory of Mind Measure for Children with Asperger Syndrome: The ATOMIC. *Journal of Autism and Developmental Disorders*, 38(2), 249-260.
- Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., et al. (2011). Theory of Mind Training in Children with Autism: A Randomized Controlled Trial. *Journal of Autism and Developmental Disorders*, 41(8), 997-1006.
- Bennett, T. A., Szatmari, P., Bryson, S., Duku, E., Vaccarella, L., & Tuff, L. (2013). Theory of Mind, language and adaptive functioning in ASD: A neuroconstructivist perspective. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 22(1), 13-19.
- Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European Psychiatry*, 20(3), 291-298.
- Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. [Review]. *Translational Psychiatry*, 3, e285.
- Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, 133(5), 1515-1525.
- Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., et al. (1996). Theory of Mind in the psychopath. *The Journal of Forensic Psychiatry*, 7(1), 15-25.
- Blair, R. J. R. (2008). Fine cuts of empathy and the amygdala: Dissociable deficits in psychopathy and autism. *The Quarterly Journal of Experimental Psychology*, 61(1), 157-170.
- Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, 9(4), 267-277.
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1), 25-31.
- Bowler, D. M. (1992). "Theory of Mind" in Asperger's Syndrome. *Journal of Child Psychology and Psychiatry*, 33(5), 877-893.

- Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of Children with Autism Spectrum Disorder on Advanced Theory of Mind Tasks. *Autism, 8*(3), 283-299.
- Brunsdon, V. E., & Happé, F. (2014). Exploring the 'fractionation' of autism at the cognitive level. *Autism, 18*(1), 17-30.
- Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The Role of Causal and Intentional Judgments in Moral Reasoning in Individuals with High Functioning Autism. *Journal of Autism and Developmental Disorders, 43*(2), 458-470.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187-192.
- Callenmark, B., Kjellin, L., Rönqvist, L., & Bölte, S. (2013). Explicit versus implicit social cognition testing in autism spectrum disorder. *Autism, 18*(6), 684-693.
- Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology, 48*(1), 257-270.
- Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839-1849.
- Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T. O. M., Meldrum, D., et al. (2007). Validation of the Social Communication Questionnaire in a Population Cohort of Children With Autism Spectrum Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(10), 1324-1332.
- Channon, S., Crawford, S., Orlowska, D., Parikh, N., & Thoma, P. (2013). Mentalising and social problem solving in adults with Asperger's syndrome. *Cognitive Neuropsychiatry, 19*(2), 149-163.
- Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology, 33*(5), 781-789.

- Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, 49(3), 507-517.
- Chong, S. F., & Choo, R. (2011). Introduction to Bootstrap. *Proceedings of Singapore Healthcare*, 20(3), 236-240.
- Chung, Y. S., Barch, D., & Strube, M. (2013). A Meta-Analysis of Mentalizing Impairments in Adults With Schizophrenia and Autism Spectrum Disorder. *Schizophrenia Bulletin*.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological science*, 24(5), 723-732.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113-126.
- de Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, 10(10), 435-441.
- Demurie, E., De Corel, M., & Roeyers, H. (2011). Empathic accuracy in adolescents with autism spectrum disorders and adolescents with attention-deficit/hyperactivity disorder. *Research in Autism Spectrum Disorders*, 5(1), 126-134.
- Devine, R. T., & Hughes, C. (2013). Silent Films and Strange Stories: Theory of Mind, Gender, and Social Experiences in Middle Childhood. *Child Development*, 84(3), 989-1003.
- Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, 55(2), 705-712.
- Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine*, 34(6), 1093-1102.

- Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13(2), 331-338.
- Dziobek, I. (2012). Comment: Towards a More Ecologically Valid Assessment of Empathy. *Emotion Review*, 4(1), 18-19.
- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: A Movie for the Assessment of Social Cognition. *Journal of Autism and Developmental Disorders*, 36(5), 623-636.
- Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H., Wolf, O., et al. (2008). Dissociation of Cognitive and Emotional Empathy in Adults with Asperger Syndrome Using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, 38(3), 464-473.
- Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E., Brammer, M., et al. (2010). Describing the Brain in Autism in Five Dimensions—Magnetic Resonance Imaging-Assisted Diagnosis of Autism Spectrum Disorder Using a Multiparameter Classification Approach. *The Journal of Neuroscience*, 30(32), 10612-10623.
- Fan, Y.-T., Chen, C., Chen, S.-C., Decety, J., & Cheng, Y. (2013). Empathic arousal and social understanding in individuals with autism: evidence from fMRI and ERP measurements. *Social Cognitive and Affective Neuroscience*. 24(5) 659
- Fernández, A. M., Dufey, M., & Kramp, U. (2011). Testing the Psychometric Properties of the Interpersonal Reactivity Index (IRI) in Chile. *European Journal of Psychological Assessment*, 27(3), 179-185.
- Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., et al. (1995). Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition*, 57(2), 109-128.
- Fodor, J. A. (1981). *The modularity of mind: An essay on faculty psychology*. Cambridge, Mass: MIT.
- Frith, U. (1991). *Autism and Asperger syndrome*. UK: Cambridge University Press.
- Frith, U. (1994). Autism and theory of mind in everyday life. *Social Development*, 3(2), 108-124.

- Frith, U., & Happé, F. (1994). Autism: beyond “theory of mind”. *Cognition*, 50(1–3), 115–132.
- Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: autism. *Trends in Neurosciences*, 14(10), 433–438.
- Gleichgerrcht, E., Torralva, T., Rattazzi, A., Marengo, V., Roca, M., & Manes, F. (2013). Selective impairment of cognitive empathy for moral judgment in adults with high functioning autism. *Social Cognitive and Affective Neuroscience*, 8(7), 780–788.
- Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The ‘Reading the Mind in Films’ Task [Child Version]: Complex Emotion and Mental State Recognition in Children with and without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 38(8), 1534–1541.
- Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The “Reading the Mind in Films” Task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience*, 1(2), 111–123.
- Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., et al. (2012). Alexithymia and the Processing of Emotional Facial Expressions (EFEs): Systematic Review, Unanswered Questions and Further Perspectives. *PLoS One*, 7(8), e42429.
- Grynberg, D., Luminet, O., Corneille, O., Grèzes, J., & Berthoz, S. (2010). Alexithymia in the interpersonal domain: A general deficit of empathy? *Personality and Individual Differences*, 49(8), 845–850.
- Hadjikhani, N., Zurcher, N. R., Rogier, O., Hippolyte, L., Lemonnier, E., Ruest, T., et al. (2014). Emotional contagion for pain is intact in autism spectrum disorders. [Original Article]. *Translational Psychiatry*, 4, e343.
- Hallerbäck, M. U., Lugnegård, T., Hjärthag, F., & Gillberg, C. (2009). The Reading the Mind in the Eyes Test: Test-retest reliability of a Swedish version. *Cognitive Neuropsychiatry*, 14(2), 127–143.

- Hallmayer, J., Cleveland, S., Torres, A., & et al. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry*, 68(11), 1095-1102.
- Happé, F. (1994). An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24(2), 129-154.
- Happé, F. (1994). *Autism: An introduction to psychological theory* London: UCL Press Limited.
- Happé, F., & Frith, U. (2006). The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5-25.
- Happé, F., & Frith, U. (2013). Annual Research Review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, 55(6), 553-577.
- Happé, F., & Ronald, A. (2008). The 'Fractionable Autism Triad': A Review of Evidence from Behavioural, Genetic, Cognitive and Neural Research. *Neuropsychology Review*, 18(4), 287-304.
- Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, 9(10), 1218-1220.
- Happé, F., & Vital, P. (2009). What aspects of autism predispose to talent? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1369-1375.
- Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child development*, 66(3), 843-855.
- Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48(2), 101-119.
- Happé, F. (1995). Understanding Minds and Metaphors: Insights from the Study of Figurative Language in Autism. *Metaphor and Symbolic Activity*, 10(4), 275-295.

- Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., & Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development. *Psychological Medicine*, 42(11), 2453-2459.
- Heavey, L., Phillips, W., Baron-Cohen, S., & Rutter, M. (2000). The Awkward Moments Test: A Naturalistic Measure of Social Understanding in Autism. *Journal of Autism and Developmental Disorders*, 30(3), 225-236.
- Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *Journal of Autism and Developmental Disorders*, 34(2), 229-235.
- Hill, E. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, 8(1), 26-32.
- Hill, E., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia*, 44(14), 2822-2835.
- Hill, E., & Frith, U. (2003). Understanding autism: insights from mind and brain. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 281-289.
- Hoekstra, R., Bartels, M., Cath, D., & Boomsma, D. (2008). Factor Structure, Reliability and Criterion Validity of the Autism-Spectrum Quotient (AQ): A Study in Dutch Population and Patient Groups. *Journal of Autism and Developmental Disorders*, 38(8), 1555-1566.
- Huws, J. C., & Jones, R. S. P. (2010). 'They just seem to live their lives in their own little world': lay perceptions of autism. *Disability & Society*, 25(3), 331-344.
- Ickes, W. (1993). Empathic Accuracy. *Journal of Personality*, 61(4), 587-610.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads (Vol. 59, pp. 730-742). US: American Psychological Association.
- Jameel, L., Vyas, K., Bellesi, G., Roberts, V., & Channon, S. (2014). Going 'Above and Beyond': Are Those High in Autistic Traits Less Pro-social? *Journal of Autism and Developmental Disorders*, 1-13.

- Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology*, 10(2), 74-81.
- Johnston, K., Madden, A., Bramham, J., & Russell, A. (2011). Response Inhibition in Adults with Autism Spectrum Disorder Compared to Attention Deficit/Hyperactivity Disorder. *Journal of Autism and Developmental Disorders*, 41(7), 903-912.
- Johnston, L., Miles, L., & McKinlay, A. (2008). A critical review of the eyes test as a measure of social-cognitive impairment. *Australian Journal of Psychology*, 60(3), 135-141.
- Jolliffe, T., & Baron-Cohen, S. (1999). The Strange Stories Test: A Replication with High-Functioning Adults with Autism or Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 29(5), 395-406.
- Jones, A. P., Happé, F. G. E., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 51(11), 1188-1197.
- Jones, C. R. G., Pickles, A., Falcato, M., Marsden, A. J. S., Happé, F., Scott, S. K., et al. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 52(3), 275-285.
- Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, 65(8), 946-954.
- Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience*, 4(2), 135-152.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2(3), 217-250.
- Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, 45(1), 56-63.
- Kasari, C., & Patterson, S. (2012). Interventions Addressing Social Impairment in Autism. *Current Psychiatry Reports*, 14(6), 713-725.

- Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences*, 16(11), 559-572.
- Klein, A. M., Zwickel, J., Prinz, W., & Frith, U. (2009). Animated triangles: An eye tracking investigation. *The Quarterly Journal of Experimental Psychology*, 62(6), 1189-1197.
- Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 345-360.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59(9), 809-816.
- Klin, A., Saulnier, C., Sparrow, S., Cicchetti, D., Volkmar, F., & Lord, C. (2007). Social and Communication Abilities and Disabilities in Higher Functioning Individuals with Autism Spectrum Disorders: The Vineland and the ADOS. *Journal of Autism and Developmental Disorders*, 37(4), 748-759.
- Kylliäinen, A., Wallace, S., Coutanche, M. N., Leppänen, J. M., Cusack, J., Bailey, A. J., et al. (2012). Affective-motivational brain responses to direct gaze in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 53(7), 790-797.
- Lahera, G., Boada, L., Pousa, E., Mirapeix, I., Morón-Nozaleda, G., Marinas, L., et al. (2014). Movie for the Assessment of Social Cognition (MASC): Spanish Validation. *Journal of Autism and Developmental Disorders*, 44, 1886-1896..
- Lang, R., Regester, A., Lauderdale, S., Ashbaugh, K., & Haring, A. (2010). Treatment of anxiety in autism spectrum disorders using cognitive behaviour therapy: A systematic review. *Developmental Neurorehabilitation*, 13(1), 53-63.
- Leekam, S. R., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition*, 40(3), 203-218.
- Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind.". *Psychological Review*, 94(4), 412-426.

- Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology*, 6(4), 315-324.
- Leslie, A. M., & Happé, F. (1989). Autism and ostensive communication: The relevance of metarepresentation. *Development and Psychopathology*, 1(03), 205-212.
- Lind, S., & Bowler, D. (2009). Language and Theory of Mind in Autism Spectrum Disorder: The Relationship Between Complement Syntax and False Belief Task Performance. *Journal of Autism and Developmental Disorders*, 39(6), 929-937.
- Lindner, J. L., & Rosén, L. A. (2006). Decoding of Emotion through Facial Expression, Prosody and Verbal Content in Children and Adolescents with Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, 36(6), 769-777.
- Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013). Dissecting empathy: high levels of psychopathic and autistic traits are characterised by difficulties in different social information processing domains. [Original Research]. *Frontiers in Human Neuroscience*, 7.
- Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One*, 2(9), e883
- Lord, C., Petkova, E., Hus, V., & et al. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry*, 69(3), 306-313.
- Lord, C., Rutter, M., & Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659-685.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, 19(2), 185-212.
- Magnée, M. J. C. M., De Gelder, B., Van Engeland, H., & Kemner, C. (2007). Facial electromyographic responses to emotional information from faces and voices in

- individuals with pervasive developmental disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 48(11), 1122-1130.
- Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *American Psychological Association*. 68, 854-869.
- Mathersul, D., McDonald, S., & Rushby, J. A. (2013a). Automatic facial responses to affective stimuli in high-functioning adults with autism spectrum disorder. *Physiology & Behavior*, 109(0), 14-22.
- Mathersul, D., McDonald, S., & Rushby, J. A. (2013b). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology*, 35(6), 655-668.
- Matson, J., Hattier, M., & Williams, L. (2012). How Does Relaxing the Algorithm for Autism Affect DSM-V Prevalence Rates? *Journal of Autism and Developmental Disorders*, 42(8), 1549-1556.
- Mazefsky, C. A., Kao, J., & Oswald, D. P. (2011). Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5(1), 164-174.
- McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. [Review]. *Disability and Rehabilitation*, 28(24), 1529-1542.
- McDonald, S., & Flanagan, S. (2004). Social perception deficits after traumatic brain injury: Interaction between emotion recognition, mentalizing ability, and social communication. [Article]. *Neuropsychology*, 18(3), 572-579.
- McDonald, S., Flanagan, S., & Rollins, J. (2002). *The Awareness of Social Inference Test*. Sydney: Harcourt Assessment.
- McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and Specificity of Proposed DSM-5 Diagnostic Criteria for Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(4), 368-383.

- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-Analysis of the Relation Between Language Ability and False-belief Understanding. *Child Development, 78*(2), 622-646.
- Milton, D. E. M. (2012). On the ontological status of autism: the 'double empathy problem'. *Disability & Society, 27*(6), 883-887.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology, 41*(1), 49-100.
- Montag, C., Dziobek, I., Richter, I. S., Neuhaus, K., Lehmann, A., Sylla, R., et al. (2011). Different aspects of theory of mind in paranoid schizophrenia: Evidence from a video-based assessment. *Psychiatry Research, 186*(2-3), 203-209.
- Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., et al. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences, 108*(7), 2688-2692.
- Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., et al. (2006). Impaired self-awareness and theory of mind: An fMRI study of mentalizing in alexithymia. *NeuroImage, 32*(3), 1472-1482.
- Moriguchi, Y., Ohnishi, T., Mori, T., Matsuda, H., & Komaki, G. E. N. (2007). Changes of brain activity in the neural substrates for theory of mind during childhood and adolescence. *Psychiatry and Clinical Neurosciences, 61*(4), 355-363.
- Muller, F., Simion, A., Reviriego, E., Galera, C., Mazaux, J.-M., Barat, M., et al. (2010). Exploring theory of mind after severe traumatic brain injury. *Cortex, 46*(9), 1088-1099.
- Mundy, P., & Sigman, M. (1989). The theoretical implications of joint-attention deficits in autism. *Development and Psychopathology, 1*(03), 173-183.
- Murphy, D. G. M., Beecham, J., Craig, M., & Ecker, C. (2011). Autism in adults. New biological findings and their translational implications to the cost of clinical services. *Brain Research, 1380*(0), 22-33.

- Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, 308(5719), 255-258.
- Palmen, A., Didden, R., & Lang, R. (2012). A systematic review of behavioral intervention research on adaptive skill building in high-functioning young adults with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 6(2), 602-617.
- Park, S. H., Goo, J. M., & Jo, C.-H. (2004). Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean Journal of Radiology*, 5(1), 11-18.
- Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, 43(4), 974-990.
- Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology*, 46(2), 530-544.
- Pellicano, E. (2012). The Development of Executive Function in Autism. *Autism Research and Treatment*, 2012, 8.
- Peterson, E., & Miller, S. (2012). The Eyes Test as a measure of individual differences: How much of the variance reflects verbal IQ? [Original Research]. *Frontiers in Psychology*, 3.
- Philip, R. C. M., Dauvermann, M. R., Whalley, H. C., Baynham, K., Lawrie, S. M., & Stanfield, A. C. (2012). A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, 36(2), 901-942.
- Pinkham, A. E., Penn, D. L., Green, M. F., Buck, B., Healey, K., & Harvey, P. D. (2013). The Social Cognition Psychometric Evaluation Study: Results of the Expert Survey and RAND Panel. *Schizophrenia Bulletin*. 40(4) 813-823.
- Ponnet, K., Buysse, A., Roeyers, H., & Clercq, A. (2008). Mind-Reading in Young Adults with ASD: Does Structure Matter? *Journal of Autism and Developmental Disorders*, 38(5), 905-918.
- Ponnet, K., Buysse, A., Roeyers, H., & Corte, K. (2005). Empathic Accuracy in Adults with a Pervasive Developmental Disorder During an Unstructured Conversation with a

- Typically Developing Stranger. *Journal of Autism and Developmental Disorders*, 35(5), 585-600.
- Ponnet, K. S., Roeyers, H., Buysse, A., De Clercq, A., & Van Der Heyden, E. (2004). Advanced Mind-Reading in Adults with Asperger Syndrome. *Autism*, 8(3), 249-266.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(04), 515-526.
- Ricketts, J., Jones, C. G., Happé, F., & Charman, T. (2013). Reading Comprehension in Autism Spectrum Disorders: The Role of Oral Language and Social Functioning. *Journal of Autism and Developmental Disorders*, 43(4), 807-816.
- Robinson, E. B., Lichtenstein, P., Anckarsäter, H., Happé, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences*, 110(13), 5258-5262.
- Roeyers, H., Buysse, A., Ponnet, K., & Pichal, B. (2001). Advancing Advanced Mind-reading Tests: Empathic Accuracy in Adults with a Pervasive Developmental Disorder. *Journal of Child Psychology and Psychiatry*, 42(2), 271-278.
- Roeyers, H., & Demurie, E. (2010). How impaired is mind-reading in high-functioning adolescents and adults with autism? *European Journal of Developmental Psychology*, 7(1), 123-134.
- Rogers, K., Dziobek, I., Hassenstab, J., Wolf, O., & Convit, A. (2007). Who Cares? Revisiting Empathy in Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 37(4), 709-715.
- Russell, J. (1997). *Autism as an Executive Disorder* Oxford: Oxford University Press.
- Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the Mind in the Voice: A Study with Normal Adults and Adults with Asperger Syndrome and High Functioning Autism. *Journal of Autism and Developmental Disorders*, 32(3), 189-194.

- Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification. *Autism Research*, 6(2), 134-143.
- Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 54(6), 628-635.
- Scholl, B. J., & Leslie, A. M. (1999). Modularity, Development and 'Theory of Mind'. *Mind & Language*, 14(1), 131-153.
- Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind Eyes: An Absence of Spontaneous Theory of Mind in Asperger Syndrome. *Science*, 325(5942), 883-885.
- Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. *Social Neuroscience*, 3(2), 97-112.
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(8), 921-929.
- Siu, A. M. H., & Shek, D. T. L. (2005). Validation of the Interpersonal Reactivity Index in a Chinese Context. *Research on Social Work Practice*, 15(2), 118-126.
- Smith, A. (2006). Cognitive Empathy and Emotional Empathy in Human Behaviour and Evolution. *Psychological Record*, 56(1), 3-21.
- Smith, A. (2009a). Emotional Empathy in Autism Spectrum Conditions: Weak, Intact, or Heightened? *Journal of Autism and Developmental Disorders*, 39(12), 1747-1748.
- Smith, A. (2009b). The empathy imbalance hypothesis of autism: a theoretical approach to cognitive and emotional empathy in autistic development. *Psychological Record*, 59(2), 489-510.
- Social, Care Local Government and Care Partnership Directorate. (2014). *THINK AUTISM Fulfilling and Rewarding Lives, the Strategy for Adults with Autism in England: an update*. London. Department of Health.

- Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, 43(5), 433-443.
- Sparks, A., McDonald, S., Lino, B., O'Donnelle, M., & Green, M. J. (2010). Social cognition, empathy and functional outcome in schizophrenia. *Schizophrenia Research*, 122(1-3), 172-178.
- Spek, A., Scholte, E., & Berckelaer-Onnes, I. (2010). Theory of Mind in Adults with HFA and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 40(3), 280-289.
- Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language*, 17(1-2), 3-23.
- Storch, E. A., Arnold, E. B., Lewin, A. B., Nadeau, J. M., Jones, A. M., De Nadai, A. S., et al. (2013). The Effect of Cognitive-Behavioral Therapy Versus Treatment as Usual for Anxiety in Children With Autism Spectrum Disorders: A Randomized, Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(2), 132-142.e132.
- Storch, E. A., Ehrenreich May, J., Wood, J. J., Jones, A. M., De Nadai, A. S., Lewin, A. B., et al. (2012). Multiple informant agreement on the anxiety disorders interview schedule in youth with autism spectrum disorders. *Journal of Child and Adolescent Psychopharmacology*, 22(4), 292-299.
- Swart, M., Kortekaas, R., & Aleman, A. (2009). Dealing with feelings: characterization of trait alexithymia on emotion regulation strategies and cognitive-emotional processing. *PLoS One*, 4(6), e5751.
- Sze, K. M., & Wood, J. J. (2008). Enhancing CBT for the Treatment of Autism Spectrum Disorders and Concurrent Anxiety. *Behavioural and Cognitive Psychotherapy*, 36(Special Issue 04), 403-409.
- Tantam, D. (2014). Adults with ASD. *Current Developmental Disorders Reports*, 1(1), 1-7.
- Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (2003). The 20-Item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychosomatic Research*, 55(3), 277-283.

- Taylor, G. J., Michael Bagby, R., & Parker, J. D. A. (1991). The Alexithymia Construct: A Potential Paradigm for Psychosomatic Medicine. *Psychosomatics*, 32(2), 153-164.
- Travers, B. G., Adluru, N., Ennis, C., Tromp, D. P. M., Destiche, D., Doran, S., et al. (2012). Diffusion Tensor Imaging in Autism Spectrum Disorder: A Review. *Autism Research*, 5(5), 289-313.
- Trikalinos, T. A., Karvouni, A., Zintzaras, E., Ylisaukko-oja, T., Peltonen, L., Jarvela, I., et al. (2005). A heterogeneity-based genome search meta-analysis for autism-spectrum disorders. *Molecular Psychiatry*, 11(1), 29-36.
- Tsai, L., & Ghaziuddin, M. (2014). DSM-5 ASD Moves Forward into the Past. *Journal of Autism and Developmental Disorders*, 44(2), 321-330.
- Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., et al. (2012). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry*, 18(4), 326-354.
- Vermeulen, P. (2014). Context Blindness in Autism Spectrum Disorder: Not Using the Forest to See the Trees as Trees. *Focus on Autism and Other Developmental Disabilities*. 1-11.
- Volden, J., & Johnston, J. (1999). Cognitive Scripts in Autistic Children and Adolescents. *Journal of Autism and Developmental Disorders*, 29(3), 203-211.
- Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A Cross-Cultural Comparison. *Journal of Autism and Developmental Disorders*, 36(2), 263-270.
- Walter, H. (2012). Social Cognitive Neuroscience of Empathy: Concepts, Circuits, and Genes. *Emotion Review*, 4(1), 9-17.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (WAIS) (3rd UK ed.)*. London: Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scales of Intelligence*. San Antonio, TX: Harcourt Assessment.

- Wechsler, D. (2008). WAIS-IV administration and scoring manual. San Antonio, TX: Psychological Corporation.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, 72(3), 655-684.
- White, S. (2013). The Triple I Hypothesis: Taking Another('s) Perspective on Executive Dysfunction in Autism. *Journal of Autism and Developmental Disorders*, 43(1), 114-121.
- White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the Strange Stories: Revealing Mentalizing Impairments in Autism. *Child Development*, 80(4), 1097-1117.
- White, S., Ollendick, T., Albano, A., Oswald, D., Johnson, C., Southam-Gerow, M., et al. (2013). Randomized Controlled Trial: Multimodal Anxiety and Social Skill Intervention for Adolescents with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 43(2), 382-394.
- White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with autism. *Autism Research*, 4(2), 149-154.
- White, S. W., Oswald, D., Ollendick, T., & Scahill, L. (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review*, 29(3), 216-229.
- Wildt, A. R., & Ahtola, O. T.. (1978). *Analysis of covariance. Quantitative Applications in the Social Sciences series #12*. Thousand Oaks, CA: Sage Publications.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103-128.
- Wimmer, H., & Weichbold, V. (1994). Children's theory of mind: Fodor's heuristics examined. *Cognition*, 53(1), 45-57.
- Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine*, 11(01), 115-129.

- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9(1), 11-29.
- Wing, L., Gould, J., & Gillberg, C. (2011). Autism spectrum disorders in the DSM-V: Better or worse than the DSM-IV? *Research in Developmental Disabilities*, 32(2), 768-773.
- World Health Organization.(1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines* (Vol. 1). Geneva: World Health Organization.
- Yang, D. J., & Baillargeon, R. (2013). Brief Report: Difficulty in Understanding Social Acting (But Not False Beliefs) Mediates the Link Between Autistic Traits and Ingroup Relationships. *Journal of Autism and Developmental Disorders*, 43(9), 2199-2206.
- Yildirim, E. A., Kasar, M., Guduk, M., Ates, E., Kucukparlak, I., & Ozalmete, E. O. (2011). Investigation of the reliability of the "reading the mind in the eyes test" in a Turkish population. *Turk Psikiyatri Derg*, 22(3), 177-186.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). It Takes Two: The Interpersonal Nature of Empathic Accuracy. *Psychological Science*, 19(4), 399-404.
- Zalla, T., Barlassina, L., Buon, M., & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition*, 121(1), 115-126.
- Zalla, T., & Leboyer, M. (2011). Judgment of Intentionality and Moral Evaluation in Individuals with High Functioning Autism. *Review of Philosophy and Psychology*, 2(4), 681-698.
- Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., et al. (2009). Clinical Assessment and Management of Toddlers With Suspected Autism Spectrum Disorder: Insights From Studies of High-Risk Infants. *Pediatrics*, 123(5), 1383-1391.

6 Appendices

6.1 SSFT Experimental Scripts

Direction: Any shots of dialogue are filmed directly from a first person perspective of the other character with a few seconds extra. All other shots are filmed from a third person perspective.

Before participants start the test researcher reads/they can read:

You are about to watch a series of short clips where two characters will be interacting with one another. Following each clip you will be asked to answer a number of questions regarding what you have seen. For all of the clips, the characters will remain the same; one man called Max and one woman called Alice. However, it is important that each clip be viewed as independent from each other and that character's actions and utterances in previous clips do not influence decisions about character's behaviour in future clips. Do you have any questions?

Qs for each clip =

- 1) 'Why did X say that?
- 2) 'If you were X, what would you have said in that situation.
- 3) Memory based question that will be decided once filmed.

Forget

SCENE 1. HALLWAY/OUTSIDE - DAY

Alice and Max stand in a hallway. Max puts on a coat. Alice holds the door open and kisses him goodbye.

ALICE

Mark lives at number 52... see you later.

CUT. Max is outside at a door.

MAX
(quite happy)
Hey Mark I'm outside your house...number 25.

SCENE 2. KITCHEN - NIGHT

Max and Alice are laying the Table for dinner. Max is dressed in a shirt and Alice is in a dress.

MAX
(neutrally)
Rob's coming tonight for dinner. (seriously)
Did you bring his present?

ALICE
(worried)
Oh No!!!

Or

(confused)

What present?

SCENE 3. KITCHEN - DAY

Max is preparing food. Alice enters with coat on.

MAX
OK I need potatoes for the dinner tonight and whatever you want for pudding...(gives Alice some change) that should be enough. Thanks.

Max is chopping veg. Cut to: Alice walks in holding cake. Max holds his hands up, confused.

ALICE
(confused)
What? I got the cake for pudding.

Pretend

SCENE 4. OFFICE - DAY

Alice and Max pretend to row across the office on chairs coming from main door.
Alice is in front and Max behind.

MAX
(smiling and looking
happy)
And one... And two... And.

ALICE
(turning to max,
slightly annoyed but
still smiling)
Hey, you splashed me.

SCENE 5. LOUNGE - DAY

Max and Alice play a computer game, which ends. She celebrates by throwing her
hands in the air.

ALICE
(ecstatic)
Yes!

Alice pretends to shoot Max with her controller. Max holds his hands to his chest.

MAX
(worried)
I'm bleeding!

SCENE 6. OUTSIDE - DAY

Max and Alice are walking through the park. Max spots the children's playground and runs towards the swings and stops by them.

MAX
(cute)
Please push me mummy!

Appearance reality

SCENE 7. BUS STOP - DAY

Alice waits at the bus stop. Max walks up to the bus stop in a superman t-shirt.

MAX
You catching the bus into town too?

ALICE
(smiling slightly)
Why are you here Max, your superpowers
not working today?

SCENE 8. OUTSIDE - DAY

Alice and Max standing in the kitchen tying some balloons to a chair.

MAX
(neutral but smiling
at the end)
Cool, people should be here any second.. do
you want a beer?

ALICE
(smiling)
I'll have some honey please.

SCENE 9 HALLWAY - DAY

Alice waits in living room drawing on a moustache. Max comes in dressed in dress and Alice checks him out

MAX

Ooo Just going to pop to the loo before we head off.

ALICE

(neutral)

I think the ladies are on the right.

White Lie

SCENE 9.5 OUTSIDE - DAY

Max comes out of the house to put some rubbish into the bin, and notices some flowers on the pavement. He looks around and picks them up.

CUT TO Max is sitting at the Table and the roses are in a vase on the Table. Alice walks in.

ALICE

(grinning happily)

Oh Max! You shouldn't have! How lovely!

Alice bends over and kisses Max.

MAX

I picked them especially as I knew Roses were your favourite.

SCENE 10. KITCHEN - DAY

Alice and Max are eating. Alice sees Max eating loads but Alice isn't eating anything.
Max raises his head looks expectant

MAX

It's my mum's recipe... what do you think of it?

ALICE

(half smile)

It's wonderful, but I'm not really hungry.

SCENE 11. LOUNGE - DAY

Max and Alice are sitting in living room. Alice is holding a guitar.

ALICE

(nervous and expectant)

I've been working on this for ages and I think I have finally got it. I think my songs gonna end like this...

She strums a badly played chord.

ALICE

Ooo ooo ooo yeah...

MAX

(nodding his head encouragingly and half smiling)

Well done Alice... that sounds really good.

Double Bluff

SCENE 12. KITCHEN - DAY

Max and Alice sit at the Table. There are finished plates of food on the Table and they are eating sweets. Alice picks up a packet of sweets and empties it and there is one left that rolls onto her hand. She looks up at Max. Max looks disappointed

ALICE
(slight sigh)
Oh ok...look away?

Alice picks up sweet and places in her left hand.

MAX
Can I look?

ALICE
Yep, which one?... think I can feel something
in my left hand...

Alice shakes her hand a little.

SCENE 13. LOUNGE - DAY

Max is on the computer scrolling through women's profiles on a dating website. Hearing the noise of Alice opening the door, Max looks up. He brings up a work document. Alice walks in.

ALICE
Hi honey, what you up to?

Alice leans over and kisses Max.

MAX
(casually)
Oh, just looking for a new girlfriend!

SCENE 14. OFFICE - DAY

Max is wearing a suit. His tie is a bit loose and shirt messy and he is sleeping at his desk. The phone rings. Max wakes looks startled and is apprehensive about answering.

ALICE
(inquisitively)
Bryan needed the report for the meeting at
midday, why haven't you sent it through?

MAX
(annoyed)
Ummm...because I've just been sleeping all
afternoon (in sarcastic tone).

Irony

SCENE 15. BEDROOM - DAY

An alarm clock shows 8:55am. Alice awakes and looks at the alarm clock with surprise.

Max is slowly shaving in the bathroom, taking his time. He hears a loud knock at the door.

MAX
(friendly)
Morning

ALICE
(annoyed but calm
voice)
I've got all day Max, take your time.

SCENE 16. KITCHEN - DAY

Alice is standing at the counter. There is lots of mess all over the side. Max walks in with lots of shopping bags and looks around in despair.

ALICE
(questioningly)
Oh do you need to use the kitchen?

MAX
(neutral)
No, this is just perfect.

SCENE 17. LOUNGE - DAY

Max and Alice are sitting on sofa. Alice is reading book and Max is preparing his insulin.

MAX
(neutral)
I love having to do this every day.

Mixed emotions

SCENE 18. OFFICE/LOUNGE - DAY

Alice is sitting in an office and receives a phone call. She picks it up.

ALICE
(nervous)
Yes...

(confident)
I'm sure I'll do a great job... Yes I'm confident
I can manage the whole team... Thank you so
much for offering me the job

Alice puts down the phone and smiles.

CUT TO Alice and Max are sitting on the sofa. Both are smiling and she is wearing the same suit. They clink glasses.

ALICE
(worried)
God this new position is going to be so
challenging.

SCENE 19. PARK – DAY

Max and Alice are in the park. Max is throwing ashes and is crying. CUT Max and Alice are sitting on bench, Max has photo album in his hands.

ALICE
(concerned)
You OK?

Max is *crying but smiling and looking happy at the photographs.*

MAX
It makes me so happy seeing us all together
in the photos

SCENE 20. LOUNGE - DAY

Max and Alice are sitting on the sofa. Max has his computer in front of him.

MAX
(a bit embarrassed)
Jim from work just sent me this disgusting
picture of his biking accident.

ALICE
(a bit
disgusted/apprehensive)
I'm squeamish... but let's have a look then?

SCENE 20.5 OFFICE - DAY

Max sitting at his desk at work and answers his telephone.

MAX

(looking and sounding low)

Hey Nick (pause) No, they gave the promotion to Alice, I really thought I'd got it this time....(looks up).sorry mate I'll call you later (hangs up)

Alice walks in.

Alice

(massive grin)

Hey

Max

(smiling genuinely)

Hey Alice...congratulations (they hug) I am so happy you got promoted

Persuasion

SCENE 21. LOUNGE - DAY

Alice is sitting with a suit case and checking her tickets. Max goes to get up.

MAX

(tired and run down)

I'm going to have a bath. I'm really tired.

ALICE

Oh please can you give me lift to the station
Max...

She sniffs melodramatically.

ALICE

I've got a cold starting I think.

SCENE 22. kitchen - DAY

Max and Alice sit at the Table and Max hands Alice a piece of paper. Alice gestures that she doesn't want it.

ALICE

Sorry Max, but we've practiced a million times, I don't really feel like going through your interview questions again right now.

MAX

Please Alice, If we don't practice tonight I'll forget everything in the interview.

SCENE 23. SOFA - DAY

Alice sits on sofa reading book. Max enters in coat and sits down next to her.

ALICE

You off then... say hi to everyone from me'

MAX

(ernest)

I know you find my mum difficult but if you came it would make me really happy.

SCENE 24. GARDEN - DAY

Max lights a cigarette as he stands outside. Alice looks disappointed.

MAX

(offended)

What?

ALICE

You're going to be a dad in 8 months time.

Lie

SCENE 25. KITCHEN/LIVING ROOM - DAY

Max answers the phone in the kitchen.

MAX

You want to speak to Alice? Ok let me just go
and get her.

Max walks into room covering phone with his hand and trying to hand it to Alice.
Alice glances at him, she doesn't want to take it.

MAX

(apologetically)

Oh um sorry she's not actually here right now, can I take a message?

SCENE 26. KITCHEN/PUB - DAY

Max has dinner laid for two and pot of food. He looks impatiently at his watch and
calls Alice on her mobile.

CUT TO Alice is at the bar of a pub and notices her phone ringing and rushes outside.

ALICE

(apologetically)

Sorry Max, I had loads of work to do, but I'm literally leaving work now.

SCENE 27. BEDROOM - DAY

The alarm clock goes off and Max hits the snooze button, rolling back into bed. Fade in and out. Max wakes up abruptly and gazes at the clock showing 9am. Cut. Max walks hurriedly into the office and sits next to Alice.

ALICE

What took you so long?

MAX

Traffic was terrible this morning!

Misunderstanding

SCENE 28. SHOP - DAY

A woman walks into a shop. Max and Alice are outside.

MAX

I'll just grab some fruit, won't be a second..

The woman is in the shop and Max approaches her – she is comparing two types of apple.

MAX

Excuse me, do you have any kiwis out the back?

SCENE 29. OUTSIDE HOUSE - NIGHT

Max approaches the front door at night and looking for his keys. He calls Alice but it goes straight to answer phone. Max makes an annoyed expression and hangs up. He then looks around and climbs up and opens front window.

CUT TO Alice who is walking along the street. She sees someone breaking in. She looks afraid and then angry.

ALICE

Hey you... I'm calling the police if you don't get down.

SCENE 30. KITCHEN - DAY

Alice sits at Table. The phone rings and she puts her hands free set on, talking while on the computer.

ALICE

Hi this is Alice speaking, ah ha... no we don't have double glazing fitted at the moment...'

Max walks behind with a birthday cake in his hand about to sing.

ALICE

(still facing
computer)

Look I really don't have time to talk right now.

MAX

(slightly annoyed)

Fine!

Idiom

SCENE 31. KITCHEN - DAY

Alice sits at kitchen Table eating some food. Max walks in and sits down.

MAX

(tired and sleepy)

Hey

ALICE

Get up with the chickens today Max?

SCENE 32. OUTSIDE - DAY

Max kneels next to his bike. His hands are covered in oil and is Alice standing over him.

ALICE

What Happened?

MAX
(In despair)
I'm all thumbs today.

SCENE 33. KITCHEN - DAY

Max and Alice are sitting at a Table.

MAX
I think that if John would just own up to what
he has done it would all be ok between us.

ALICE
Well Max... it takes two to tango.

Joke

SCENE 34. OFFICE - DAY

Max is sitting at his desk. Alice enters and puts cup of coffee next to him. He sips a cup of coffee and pulls a disgusted face

MAX
Alice This coffee tastes like mud.

ALICE
(smiling)
Yes, it's fresh *ground*.

SCENE 35. KITCHEN - DAY

Max and Alice sit at breakfast Table. Max is reading a newspaper. He folds it up and puts it down, rolling his eyes.

ALICE
What have the party got say about
everything that's Happéning then?

MAX
Well unfortunately Alice, politicians are like
bins. They should both be changed regularly,
and for the same reason.

SCENE 36. LOUNGE - DAY

Max is sitting on the sofa watching TV. Alice walks in and sits down.

ALICE

You watching the news?

MAX

Yeah... have you noticed that they always
start with good evening on this program and
then proceed to tell you why it isn't.

6.2 SSFT Control Scripts

SCENE 37. KITCHEN - DAY

MAX is on the PHONE. ALICE is sitting down at the Table.

MAX

Uh-huh

MAX puts his hand over the end of the PHONE and turns to ALICE.

MAX

Ok, we can pay for our line rental monthly, which will cost us 15 pound a month or we can pay 120 now for the year.

ALICE

Well we both have the money now so lets pay for the year.

Over the year £120 works out cheaper than £15 a month

SCENE 38. KITCHEN - DAY

MAX and ALICE are sitting at the Table. MAX is holding a utility bill.

MAX

Alice... We've got November's gas bill to pay. We can pay the bill now based on an average of what we've used in the last six month. Or, we can send them a reading from the meter and pay based on that.

ALICE considers the options.

ALICE

They use the last six months... Ok... let's pay using the estimate.

The estimated bill will be based on the summer months estimate, which will be cheaper than a meter reading of an autumn month.

SCENE 39. KITCHEN - DAY

ALICE and MAX sit in front of a half finished meal.

ALICE

God I'm stuffed. I don't think I can eat anymore but I don't want to waste anything.

MAX

Me too... I know, I'll make soup tomorrow.

Making soup will use up the left over food.

SCENE 40. KITCHEN - DAY

MAX and ALICE sit at the Table reading a newspaper.

ALICE

They've announced a hose pipe ban for this summer, so we won't be able to use the hose for watering the garden.

MAX

Sure, I won't.... Before we head out I'm gonna take a shower.

ALICE

Remember to put the bucket in the shower.

They can use the shower water to water the garden plants.

SCENE 41. LOUNGE - DAY

ALICE sits on the sofa. MAX walks in with a shopping bag in his hand.

ALICE

Hey... did you pick up that second hand climbing harness?

MAX

No... Even though I'm really short of money and it cost twice as much, I went to the shop and brought a new one.

It is more important for a climbing harness to be safe than for it to cost less.

SCENE 42. KITCHEN - DAY

ALICE and MAX sit at the Table. Alice is on a computer.

ALICE

I have a choice of two anti-malarial pills for my beach holiday. They cost the same, but they have different side effects. One makes you sleepy and the other makes you sensitive to the sun. What do you think?

MAX

I think you should buy the pill that makes you sleepy.

It would be hard to enjoy a beach holiday in Thailand while being photosensitive.

SCENE 43. SOFA - DAY

MAX sitting on sofa and Alice comes and joins. Gestures to plant on coffee Table.

MAX

I got this new tropical plant today.

ALICE

Great.... You could put it in the south facing bathroom upstairs.

The plant would like humid and light conditions.

SCENE 44. LOUNGE - DAY

MAX and ALICE are sitting with their legs up looking at each other.

ALICE

Do you need me to water anything while you're away for two weeks?

MAX

No, I only have one plant and it's a cactus.

Desert plants do not need to be watered regularly.

SCENE 45. KITCHEN - DAY

ALICE and MAX sit at the Table. ALICE is reading a newspaper. She puts it down and turns to MAX.

ALICE

They are saying in the news that petrol prices are going to reach record highs this summer.

MAX

Maybe we should catch the train to Newcastle to visit your family then?

Petrol prices are affecting the cost of driving.

SCENE 46. KITCHEN - DAY

ALICE and MAX are sitting at the Table. MAX is reading a newspaper, which he puts down.

MAX

(shocked)

Oh my god...there has been an outbreak of mad cow disease in England again.

ALICE

Maybe we shouldn't eat those beef burgers then.

The beef burgers could be carrying the disease.

SCENE 47. KITCHEN

ALICE and MAX are sitting at the Table. MAX is reading a newspaper, which he puts down.

MAX

Apparently al of the soya and wheat crops have been severely affected by droughts this year.

ALICE

(concerned)

God..I hope Farmer Tucket can afford to keep his heard of cows.

The drought will increase the cost of soy and wheat. Famers who feed their cows on soy and wheat often have to kill their cattle if the feed becomes too expensive.

6.3 SSFT Project Brief for Actors.

Thank you for your interest in the project,

I would like to just briefly remind you of the project details.

A defining feature of Autism Spectrum Disorders (ASDs) is an inability to infer and/or represent complex mental states of other people. The integration of non-verbal cues e.g. facial expression, body language, and verbal utterances within specific social contexts is something we all do unconsciously to decipher the meaning of what people are communicating, but people with autism can find this challenging. Adults who have an ASD but also have average intellectual abilities are generally able to understanding literal verbal utterances where the context and the person's intentions are concordant with what is said. However, when someone communicates using figurative language e.g. sarcasm or pretence, one has to ignore the literal interpretation of the verbal information and rely on other cues e.g. tone of voice, context etc to decipher the speaker's intention. We are trying to design a task that taps into this difficulty in ASD and each clip we are filming will involve one character uttering a piece of figurative language.

As I mentioned on the phone, the most important feature of this work is that the acting is naturalistic and in some ways the project is reliant on this. The clips have been designed to be intentionally short to account for memory difficulties in the sample. Portraying the linguistic concept in this single utterance is therefore all the more important and potentially challenging.

Below find a short brief about the characters:

The characters are both approximately 30 years old and working professionals. They are in a medium to long-term relationship and work with another, but are not living with each other.

They both live in London and work in the service industry. Both their characters find their work satisfying although on a day-to-day basis they can find it boring.

Both are confident individuals who are happy in the relationship, but sometimes act in ways that suggest they are not committed to each other. However, they are both caring people and supportive of one another when it matters.

Your audition time is XXX Please be on time and text my mobile XXX to let me know that you have arrived so that I can let you into the building.

Our address is

4 Windsor Walk
Denmark Hill
London SE5 8AF

The closest station is Denmark Hill which is literally 2mins walk from the office and regular trains depart from Victoria, Blackfriars and London Bridge.

Alternatively, the number 40, 176 and 68 all stop outside the Maudsley Hospital, which is a 5-10minute walk away.

6.4 Pilot Information Sheet and Consent Form

Information about the research

PNM/10/11-22 Strange Situations



Researchers: Kim Murray, Charlotte Wormald, Dr Ailsa Russell

We would like to invite you to help us with our research. Please take time to read this leaflet carefully. You can ask us if there is anything that is not clear or if you would like further information. You will be given another copy of this information sheet and a copy of your signed consent form, should you agree to take part in the study.

What is the study about?

We are interested in the way adults perceive social, everyday situations. We hope that by studying this in adults who do not have a diagnosis of an Autism Spectrum Disorder (ASD), we will be able to understand a bit more about why some people with ASD find this more difficult.

How do I qualify for your study?

You must not have received a diagnosis for an autistic spectrum disorder or a psychiatric disorder. You must be at least 18 years old and be able to read and understand English.

Do I have to take part?

No, it is up to you to decide. You are free to leave our study at any time without giving a reason. If you decide to withdraw from the study we will not use the data collected up to your withdrawal.

What will happen to me if I take part?

1. First of all after reading this sheet, please fill out the questionnaire and email or send it back to us as soon as possible. Please use the enclosed stamped address envelope.
2. When the research assistant receives your completed questionnaire, they will contact you to let you know if we would like you to attend for a further session. This session will take place at the Institute of Psychiatry, King's College London and involve being asked to read some very short stories and answer some questions about them. Finally, you will be asked to watch some very short film clips and answer some questions about them. The whole meeting will take about 90 minutes.

Payment

You will not be paid for taking part in this study. If you approve we can add all participant's names into an anonymous prize draw to win a book voucher. The research assistant will let you know if you have won.

Recording of sessions

We will digitally audio-record the interviews with the research assistant during the assessment and the recordings will be anonymised. The recordings will be kept safe in computerised form at the Institute of Psychiatry, KCL for four years. After that time, all files of the recordings will be permanently deleted.

Will there be any disadvantages to taking part?

If you join the study, you will need to come to the Institute of Psychiatry for one assessment session lasting ninety minutes. A potential disadvantage therefore could be the time involved in attending the session.

What are possible benefits of taking part?

Our study should lead to a better understanding of the processes that underlie the way people perceive and interpret social situations. By taking part you are helping us to do this and also helping others who have difficulty in this area.

What if there is a problem or a complaint?

If this study has harmed you in any way you can contact King's College London using the details at the bottom of the information sheet for further advice and information.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice. Your data will be anonymised and kept strictly confidential and held in a locked filing cabinet or on a secure computer. People on our research team will only see your records if they need to for the research. The data and recordings that we collect will be securely stored for four years after the end of the trial, for your protection and to allow good clinical practice (GCP). After this time, data will be securely destroyed.

Your name, address and telephone number will be held securely on Institute of Psychiatry computer drives and will be password protected. This will be used only to monitor recruitment. You will not be named in any published results from our study.

What will happen to the results of the research study?

Our results may be presented at national and international conferences and published in psychological/medical journals. The results will not say who took part or give any details that could lead to anyone being identified or recognised.

Who is organising and funding the research?

The study is organised by the Institute of Psychiatry.

Who has reviewed the study?

The study (PNM/10/11-22 **Strange Situations**) has been approved by the King's College London Psychiatry, Nursing & Midwifery Research Ethics Subcommittee (PNM RESC).

Where can I get more information?

If you have any further questions, you may discuss these by contacting:

Mr Kim Murray –

Department of Forensic and Neurodevelopmental Sciences

Institute of Psychiatry

London,

SE5 8AF Tel: 0207 848 0820

CONSENT FORM FOR PARTICIPANTS IN RESEARCH STUDIES

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.



Title of Study: Strange Situations

King's College Research Ethics Committee Ref: PNM/10/11-22

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

Please tick or initial

- I understand that if I decide at any time during the research that I no longer wish to participate in this project, I can notify the researchers involved and withdraw from it immediately without giving any reason. Furthermore, I understand that I will be able to withdraw my data up to the point of publication. ☐
- I consent to the processing of my personal information for the purposes explained to me. I understand that such information will be handled in accordance with the terms of the Data Protection Act 1998. ☐
- I consent to having my assessment audio-recorded. ☐
- I am happy to be contacted by researchers about future research. ☐
- The information I have submitted may be published as a report. If you wish to receive a copy of this report, please tick this box. Please note that confidentiality and anonymity will be maintained and it will not be possible to identify you from any publications. ☐

Participant's Statement:

I _____

agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed

Date

Investigator's Statement:

I _____

confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the participant.

Signed

Date

6.5 The Autism Quotient

The Adult Autism Spectrum Quotient (AQ)
Ages 16+

SPECIMEN, FOR RESEARCH USE ONLY.

For full details, please see:

S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin and E. Clubley, (2001)
[The Autism Spectrum Quotient \(AQ\) : Evidence from Asperger Syndrome/High Functioning Autism, Males and Females, Scientists and Mathematicians](#)
Journal of Autism and Developmental Disorders 31:5-17

Name:..... Sex:.....

Date of birth:..... Today's Date:.....

How to fill out the questionnaire

Below are a list of statements. Please read each statement very carefully and rate how strongly you agree or disagree with it by circling your answer.

DO NOT MISS ANY STATEMENT OUT.

Examples

| | | | | |
|---|------------------|----------------|-------------------|---------------------|
| E1. I am willing to take risks. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| E2. I like playing board games. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| E3. I find learning to play musical instruments easy. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| E4. I am fascinated by other cultures. | definitely agree | slightly agree | slightly disagree | definitely disagree |

| | | | | |
|---|------------------|----------------|-------------------|---------------------|
| 1. I prefer to do things with others rather than on my own. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 2. I prefer to do things the same way over and over again. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 3. If I try to imagine something, I find it very easy to create a picture in my mind. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 4. I frequently get so strongly absorbed in one thing that I lose sight of other things. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 5. I often notice small sounds when others do not. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 6. I usually notice car number plates or similar strings of information. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 7. Other people frequently tell me that what I've said is impolite, even though I think it is polite. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 8. When I'm reading a story, I can easily imagine what the characters might look like. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 9. I am fascinated by dates. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 10. In a social group, I can easily keep track of several different people's conversations. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 11. I find social situations easy. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 12. I tend to notice details that others do not. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 13. I would rather go to a library than a party. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 14. I find making up stories easy. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 15. I find myself drawn more strongly to people than to things. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 16. I tend to have very strong interests which I get upset about if I can't pursue. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 17. I enjoy social chit-chat. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 18. When I talk, it isn't always easy for others to | definitely agree | slightly agree | slightly disagree | definitely disagree |

| | | | | |
|---|------------------|----------------|-------------------|---------------------|
| get a word in edgeways. | agree | agree | disagree | disagree |
| 19. I am fascinated by numbers. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 20. When I'm reading a story, I find it difficult to work out the characters' intentions. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 21. I don't particularly enjoy reading fiction. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 22. I find it hard to make new friends. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 23. I notice patterns in things all the time. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 24. I would rather go to the theatre than a museum. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 25. It does not upset me if my daily routine is disturbed. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 26. I frequently find that I don't know how to keep a conversation going. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 27. I find it easy to "read between the lines" when someone is talking to me. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 28. I usually concentrate more on the whole picture, rather than the small details. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 29. I am not very good at remembering phone numbers. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 30. I don't usually notice small changes in a situation, or a person's appearance. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 31. I know how to tell if someone listening to me is getting bored. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 32. I find it easy to do more than one thing at once. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 33. When I talk on the phone, I'm not sure when it's my turn to speak. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 34. I enjoy doing things spontaneously. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 35. I am often the last to understand the point of a joke. | definitely agree | slightly agree | slightly disagree | definitely disagree |

| | | | | |
|--|------------------|----------------|-------------------|---------------------|
| 36. I find it easy to work out what someone is thinking or feeling just by looking at their face. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 37. If there is an interruption, I can switch back to what I was doing very quickly. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 38. I am good at social chit-chat. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 39. People often tell me that I keep going on and on about the same thing. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 40. When I was young, I used to enjoy playing games involving pretending with other children. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 41. I like to collect information about categories of things (e.g. types of car, types of bird, types of train, types of plant, etc.). | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 42. I find it difficult to imagine what it would be like to be someone else. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 43. I like to plan any activities I participate in carefully. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 44. I enjoy social occasions. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 45. I find it difficult to work out people's intentions. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 46. New situations make me anxious. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 47. I enjoy meeting new people. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 48. I am a good diplomat. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 49. I am not very good at remembering people's date of birth. | definitely agree | slightly agree | slightly disagree | definitely disagree |
| 50. I find it very easy to play games with children that involve pretending. | definitely agree | slightly agree | slightly disagree | definitely disagree |

Developed by:
The Autism Research Centre
University of Cambrid

6.6 SSFT Pilot Instructions

Order A

You are about to watch a series of short video clips. In each clip there will usually be the same two characters; a man called Max and a woman called Alice. Max and Alice are boyfriend and girlfriend, and live and work together. Each video clip is self-contained, and what happens in one clip doesn't relate to what happens in other clips; there isn't a story running from one clip to another.

I will show you each clip only once and after each clip has finished I will ask you three questions about what you have just seen. I will write down your answers and also audio record them in case I miss anything. Some of the questions do not have right or wrong answers – I'm interested in what you think is going on in the clips, so please just answer the questions as best you can.

Do you have any questions?

[If participant asks questions regarding the content of the material e.g. "How long have they been in a relationship?" Respond with:

Please give your answers based just on what you see and the information I gave you at the beginning. Would you like me to re-read the instructions?]

[During testing :

If the participant asks in relation to the first question, "said what" or something to this effect, please respond with 'What they just said?'

If participants cannot generate an answer to the second question that incorporates direct speech so for example say 'I would make a joke back', prompt them by saying 'for example?' or 'such as?'. Make a note of the use of a prompt and if they cannot generate an answer.

Now you are going to watch a series of clips and I will ask you to answer some questions after you have viewed each clip just once.

Item 1 Forget (potatoes)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Alice buy from the shop?

Item 2 Control (Line Rental)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Who was Max talking to on the phone?

Item 3 Lie (Pub)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Where was Alice when she spoke to Max on the phone?

Item 4 Control (Soup)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max going to cook the next day?

Item 5 Misunderstanding (burglar)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Max doing when Alice called out?’

Item 6 Mixed Emotions (biking accident) – which bit?

Why did Alice say that?

If you were in Max’s situation, what would you say next??

What was the disgusting picture of?’

Item 7 Persuasion (interview)

Why did Max say that?

If you were in Alice’s situation, what would you say next??

What did Max pass across the Table to Alice?

Item 8 (Misunderstanding (Birthday Cake)

Why did Max say that?

If you were in Alice’s situation, what would you say next??

What was Max carrying?.

Item, 9 DB (sleep)

Why did Alice say that?

If you were in Max’s situation, what would you say next??

What was Alice doing before the phone rang?

Item 10 Irony (injection)

Why did Max say that?]

If you were in Alice’s situation, what would you say next??

What was Max doing to his tummy?.

Item 11 Pretend (computer game)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max holding in his hands?

Item, 12 AR (bear)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice wearing?.

Item, 13 Control (hose pipe)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice reading?.

Item 14 Joke (coffee)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What had Alice made for Max?.

Item 15 Forget (present)

Why did Alice say that?

If you were in Max's situation, what would you say next?

Where had Alice just come from?

Item 16 Irony (late)

Why did Alice say that?

If you were in Max's situation, what would you say next??
Where was Alice standing?.

Item 17 ME (Job)

Why did Max say everything that he said in the clip?
If you were in Alice's situation, what would you say next??
What was Max talking about on the phone?

Item 18 Control (climbing harness)

Why did Max say that?
If you were in Alice's situation, what would you say next??
What did Max have in his plastic bag?

Item 19 Joke (politics)

Why did Max say that?
If you were in Alice's situation, what would you say next??
What was Max reading?

Item 20 Misunderstanding (fruit)

Why did Max say that?
If you were in the lady in the shop's situation, what would you say next??
What did Max want to buy?

.

Item 21 Control (pills)

Why did Max say that?
If you were in Alice's situation, what would you say next??
What pills did Max recommend?

Item 22 White Lie (song)

Why did Max say that?

If you were in Alice's situation, what would you say next??

'What instrument was Alice playing?

Item 23 Control (train)

Why did Max say that?

If you were in Alice's situation, what would you say next??

'What form of transport did Max suggest?']

Item 24 Idiom (early riser)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice eating?.

Item 25 Lie (phone call)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max doing before the phone rang?']

Item 26 ME (urn)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice looking at when she was sitting on the bench?

Item 27 Irony (shopping)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max carrying?

Lie, 28 Forget (house)

Why did Max say that?

If you were in Mark's situation, what would you say next??

Who did Max phone??

Item 29 Pretend (playground)

Why did Max say that?

If you were in Alice's situation, what would you say next??

Where was Max sitting at the end?

Item 30 Persuasion (cold)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Where did Alice ask for a lift to?

Item 31 Control (cows)

Why did Alice say that?

If you were in Max's situation, what would you say next??

'What did Max say was affected by the drought?'

Item 32 White Lie (roses)

Why did Max say that?

If you were in Alice's situation, what would you say next??

Where did Max get the flowers?

Item 33 Persuasion (smoking)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Alice say Max was going to be?

Item 34 AR (superman)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Max holding?

Item 35 White Lie (dinner)

Why did Alice say that?

If you were in Max's situation, what would you say next??

How much of her food had Alice eaten?

Item 36 Control (cactus)

Why did Max say that?

If you were in Alice's situation, what would you say next??

How long was Max going away for?

Item 37 Idiom (bike)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max fixing?

Item 38 Control (gas bill)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Max reading?

Item 39 Lie (late)

Why did Max say that?

If you were in Alice's situation, what would you say next??

Where was Max at the beginning of the clip?

Item 40 Idiom (tango)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Max say John should do?

Item 41 AR (dress)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Alice about to put on Max?'

Item 42 Mixed Emotions (job)

Why did Alice say everything that she said in the clip?

If you were in Max's situation, what would you say next??

What did Alice and Max do with their glasses?

Item 43 Persuasion (mum)

Why did Max say that?

She wants her to come and see his mum.

If you were in Alice's situation, what would you say next??

Who does Alice find really difficult?'

Item 44 DB (sweet)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Alice have in her LEFT hand?

Item 45 Control (plant)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Where did Alice say the plant could go?

Item 46 Pretend (rowing)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice sitting on?

Item 47 DB (dating website)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max looking at on the computer at the start?

Item 48 Joke (news)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max watching on TV?

6.7 Scene Selection

| Theme | Content | Definitely remove | Max 1st | alice 1st | LR | kitchen | out/off | Max 2nd | alice 2nd | LR | kitchen | out/off |
|---------|------------------------|-------------------------|---------|-----------|----|---------|---------|---------|-----------|----|---------|---------|
| Forget | <i>House number</i> | | | | | | | 1 | | | | 1 |
| | present | aware she has forgotten | | | | | | | | | | |
| | <i>Potatoes</i> | | | 1 | | 1 | | | | | | |
| Pretend | <i>Rowing</i> | | | 1 | | | 1 | | | | | |
| | <i>Computer game</i> | | | | | | | 1 | | 1 | | |
| | <i>playground</i> | | | | | | | | | | | |
| AR | <i>Superman</i> | | | | | | | | | | | |
| | <i>Bear/Wolf</i> | | | | | | | | 1 | 1 | | |
| | <i>cross dressing</i> | | 1 | | 1 | | | | | | | |
| WL | <i>Roses</i> | Not a WL | | | | | | | | | | |
| | <i>Burnt food</i> | | | | | | | | 1 | | 1 | |
| | <i>Song</i> | | 1 | | 1 | | | | | | | |
| DB | <i>Sweet</i> | | | | | | | | 1 | | | |
| | <i>Dating website</i> | | 1 | | 1 | | | | | | | 1 |
| | <i>Sleep</i> | 5 exact | | | | | | | | | | |
| Irony | <i>bathroom</i> | | | | | | | | 1 | | | 1 |
| | <i>Mess</i> | | | | | | | | | | | |
| | <i>Injection</i> | | 1 | | 1 | | | | | | | |
| ME | <i>promotion</i> | | | | | | | | | | | |
| | urn | 2 exact | | | | | | | | | | |
| | <i>no job</i> | | | | | | | 1 | | | | 1 |
| | <i>biking accident</i> | | | 1 | 1 | | | | | | | |

| Theme | Content | Definitely remove | Max 1st | alice 1st | LR | kitchen | out/off | Max 2nd | alice 2nd | LR | kitchen | Out/Off |
|------------|-------------------|-------------------|---------|-----------|----|---------|---------|---------|-----------|----|---------|---------|
| Persuasion | Cold | | | | | | | | 1 | 1 | | |
| | interview | 2 exact | | | | | | | | | | |
| | smoking | | | 1 | | 1 | | | | | | |
| | mum | 3 exact | | | | | | | | | | |
| Lie | Phone | 4 exact | | | | | | | | | | |
| | pub | | | 1 | | | 1 | | | | | |
| | Late | | | | | | | 1 | | | | 1 |
| Mis | Fruit | new character | | | | | | | | | | |
| | Birthday cake | | | | | | | 1 | | | 1 | |
| | burglar | | | 1 | | | 1 | | | | | |
| Idiom | early riser | 5 exact | | | | | | | | | | |
| | bike | | | | | | | 1 | | | | 1 |
| | argument | | | 1 | | 1 | | | | | | |
| Joke | coffee | | | | | | | | | | | |
| | Politics | | 1 | | | 1 | | | | | | |
| | News | | | | | | | 1 | | 1 | | |
| | Social totals | | 1 | 4 | 0 | 3 | 2 | 4 | 1 | 2 | 1 | 2 |
| Control | line rental | | | 1 | | 1 | | | | | | |
| | gas bill | 4 exact | | | | | | | | | | |
| | soup | | 1 | | | 1 | | | | | | |
| | hose pipe | | | | | | | | 1 | | 1 | |
| | climboing harness | 3 exact | | | | | | | | | | |
| | pills | | | | | | | 1 | | | 1 | |
| | tropical plant | | | 1 | 1 | | | | | | | |
| | cactus | | | | | | | | | | | |
| | newcastle | | | | | | | 1 | | | 1 | |
| | cows | 2 exact | | | | | | | | | | |
| | Control totals | | 1 | 2 | 1 | 2 | 0 | 2 | 1 | 0 | 3 | 163 0 |
| | Social totals | | 1 | 4 | 0 | 3 | 2 | 4 | 1 | 2 | 1 | 2 |
| | Overall totals | | 2 | 6 | 1 | 5 | 2 | 6 | 2 | 2 | 4 | 2 |

6.8 SSFT Scoring

Scoring System

In all cases participants are awarded points for their best answer when multiple answers are given.

Mental State Scoring

0 points - no mental state words.

1point – simple mental state words regarding one character or another character's actions e.g. she was anxious, she wanted him to stop smoking OR words that imply psychological states in social context– defending, bullying, compromise, acknowledging, concerned etc.

2 points – meta-cognitive statements e.g. beliefs about beliefs OR intentions to affect another person's mental state e.g. she wanted him to feel guilty OR complex collection of mental states e.g. she felt both squeamish and curious at the same time..

Accuracy Scoring

Idiom

Why?

2 points –reference to Alice thinking that Max is also to blame for what has happened between them,; Any reference that doesn't imply that only John is at fault

Key words : blame, equal responsibility, guilty as well, fault

1 point – Simple description of events (things are complicated, it's a figure of speech); response that describes Alice's position but doesn't suggest shared responsibility (she is taking/defending John's side, she disagrees with Max, she thinks he should make more effort); mention of shared responsibility but responsibility is placed on Alice rather than Max.

0 points – irrelevant or incorrect information.

Mental state 2 points:

She wants max to know that she is also responsible for what has Happened

Max obviously feels as if he is in the wrong and alice is implying max is also in the wrong because he cooperated in some way or is partially responsible for the thing that has happened

What Next?

2 point – A response that acknowledges implication of blame and attempts to clarify, reconcile or defend self in situation

1 point - Poorly elaborated description (e.g. id get defensive), or direct speech (e.g. I disagree it's just John's fault or I agree it does take two to tango).

0 points – don't know, response that shows incomplete understanding e.g. what do you mean?; socially inappropriate e.g. nothing, or irrelevant response e.g. 'its their problem let them sort it out'. Statement that incorporates Alice as having role in situation.

Memory

1 point - mention of Max needing to own up, confess or admit to what he had done.

0 points – don't know or cant remember or incorrect recall.

Control – Soup

Why?

2 points – reference to soup being solution to dilemma of feeling full but not wanting to waste food; he can use the leftovers to make soup.

1 point – reference to facts (there is food left); states (he and Alice are full; not needing to eat anymore). Traits (being practical, kind or nice) or feelings (guilty).

0 – reference to irrelevant or incorrect factors (he had to do something, not cause an argument, soup is lighter).

Mental state 2 point:

He was suggesting a way of overcoming their dilemma of not wanting to eat and not wanting to waste anything

What Next?

2 point – response that shows acknowledgement of solution even if person doesn't agree or provides alternative solution.

1 point – no acknowledgment of solution e.g. go ahead, stating preference without acknowledgment e.g. I don't like soup.

0 points – don't know, socially inappropriate in the context of the clip e.g. 'let's just finish it', irrelevant or incorrect response e.g. don't throw it away.

Memory

1 point – Soup

0 points – don't know or can't remember or incorrect recall.

Mixed Emotions

Why?

2 points – reference to curiosity getting the better of her or overriding/co-occurring with her squeamishness or reservation.

1point - mention of just curiosity or squeamishness; mentions both curiosity and squeamishness, but where squeamishness or curiosity is thought to be feigned; facts e.g. its a gruesome picture.

0 points – irrelevant/incorrect factors or facts

Mental state 2 point:

Presumably she was inquisitive about the injury but also had a sense of reluctance and fear maybe

What Next?

2 point – acknowledgement of reluctance or mixed feelings. Response that highlights that her curiosity got the better of her squeamishness.

1 point – no acknowledgment of emotions but just showing her the picture, commenting on the picture or saying ‘have a look’. Statement that suggests squeamishness was feigned.

0 points – don’t know, socially inappropriate e.g. we can’t watch it until you beg, irrelevant response or response that doesn’t comprehend holding both emotions ‘why would you look if you are squeamish?’

Memory

1 point- mentions or describes accident or injury.

0 points – don’t know or can’t remember or incorrect recall.

Misunderstanding

2 points – reference to Alice mistaking or thinking Max was a burglar or that she thought she was being burgled.

1 points - reference to facts (someone was going in the window), state (she was surprised) or descriptions (it looked like someone was breaking into the house) or statement of Alice not knowing it was Max without articulating her misunderstanding.

0 points – factually incorrect or irrelevant answers; mentions thinking that someone was going to burgle her house.

Mental state 2 point:

She felt a sense of fear/anxiety about what she thought was someone breaking into her house so she threatened calling the police

What Next?

2 point – statement or action that resolves misunderstanding by revealing identity or explaining situation e.g. don't worry it's only me' or 'I forgot my keys' or 'It's me'

1 point – minimal statement that is partially correct e.g. I live here

0 points – don't know, inappropriate to the social context e.g. nothing, or showing annoyance with Alice for the misunderstanding, irrelevant response.

Memory

1 point – any response iterating he was trying to enter the house e.g. climbing through a window or trying to get into the house. NB if participant believes character was a burglar breaking into the house then is awarded memory point for articulating this.

0 points – don't know, can't remember or incorrect recall.

Lie

2 points - reference to her lying or that she doesn't want him to know she is in the pub.

1 point –partially correct e.g. cover up she is in the pub, reference to feelings without elaboration e.g. she feels guilty; facts (she is in a pub, she was supposed to be home for dinner); giving him information (she is going to be home soon), but which doesn't imply wanting to alter his belief.

0 points – incorrect intentions are assumed (e.g. having an affair, hide drinking habit), without mention of a lie or statement taken literally or irrelevant facts/factors.

2 point mental state answer:

She was lying about where she was cos she didn't want him to know she was in the pub

What Next?

2 point – statement that accepts information and articulates opinion regarding situation or requests more information e.g. 'Oh okay, do you know what time you're coming back?' or response that questions Alice's lie in a socially appropriate manner or makes a joke out of situation e.g. 'I can hear the fruit machines'.

1 point –minimal response that is still socially appropriate e.g. okay. dinners on the Table.

0 points – don; know, inappropriate to the social context e.g. accepts lie but is still annoyed or states utterance is a lie, or irrelevant response.

Memory

1 point – in a bar or a pub.

0 points – don't know, can't remember or incorrect recall.

Forget

2 points – any response that references forgetting, misheard or not paying attention.

1 point – partial description that misses central point of forgetting e.g. she was justifying her purchase; state (she is surprised, he was annoyed), he looked at her questioningly; facts (e.g. she bought a cake).

0 points – incorrect factors (e.g. she didn't want to buy potatoes or was only thinking of herself) , or facts (e.g. 'she couldn't find anything else for pudding').

2 point mental state answer:

she didn't know what she had forgotten

What Next?

2 point – statement that acknowledges Alice forgetting and/or clarifies original request/Max's previous statement appropriately or an appropriate joke that references forgetting e.g. of course you remembered the cake.

1 point –simple reassurance e.g. don't worry, just leave it, or positive regard for decision without clarification that something was forgotten e.g. great.

0 points – don;t know, inappropriate e.g. implying that forgetting was intentional, or irrelevant response.

Memory

1 point - cake or pudding

0 points – don't know, can't remember or incorrect recall.

Appearance Reality

Why?

2 points – reference to him making a joke about the fact he is dressed as a woman or they have switched roles and are pretending to be in character.

1 point – desire (he needs the toilet, he doesn't want to put the lipstick on), physical state (he is wearing women's clothing), trait (he is a comedian).

0 points – irrelevant or incorrect answers.

2 point mental state answer:

He was joking about wanting to go to the toilet dressed as a lady.

What Next?

2 point – statement that acknowledges or carries on the joke of being dressed in women's clothing/'being a woman' and that he needs the toilet e.g. 'will you do it sitting down now then?'

1 point – minimal response of what would do/say but still recognizes joke e.g. laugh, state it's funny.

0 points – don't know, inappropriate (e.g. simple acceptance of request) or takes statement literally with no appreciation of joke/role play, or irrelevant answer.

Memory

1 point – lipstick

0 points – don't know, can't remember or incorrect recall.

Irony

2 points – any mention of him being sarcastic; he is saying the opposite to what he feels /expressing the contrary.

1 point – reference to being unwell (diabetes), a physical act (he just injected himself) or statements that highlight his thoughts or intentions behind making his comment without mention of irony or sarcasm (he doesn't really like doing it, it's a chore/drag, sharing his thoughts, make light of the situation)

0 points – incorrect factors e.g. he enjoys it, or facts e.g. he is taking drugs, or irrelevant answers.

2 point mental state answer:

He was being sarcastic because he doesn't really like it and wants his girlfriend to feel sorry for him

What Next?

2 point – statement that expresses sympathy or understanding at how hard it is or trying to make him feel better while reminding him of the importance of doing it.

1 point –simple reminder of importance of using the injection e.g. 'it's for your own health' or that situation could be worse 'it's keeping you alive'. Minimal statement of sympathy that shows comprehension of sarcasm (e.g. sorry).

0 points – don't know, inappropriate e.g. nothing or patronizing or jovial remark or dismisses of character's emotion e.g. 'get on with it', or irrelevant answer.

Memory

1 point - injecting himself, using insulin, taking medication or related to diabetes (a sugar boost). NB If they believed he was taking drugs then taking drugs is awarded point.

0 points – don't know, can't remember or incorrect recall.

Control / Line rental

Why?

2 points – statements that mention it is cheaper, they save £30 or it's a better deal.

1 point – expressing her opinion, reference to having the money now and/or it being more convenient/sensible /making financial sense or not wanting to worry about the money without mention of it being cheaper.

0 points – incorrect or inappropriate response.

2 point mental state answer:

She wants Max to know that she would prefer to save the money and pay up front

What Next?

2 point – statement that acknowledges idea, shows agreement with option chosen and/or with proviso e.g. checking finances.

1 point minimal description of what would say or do.

0 points – don't know, socially inappropriate e.g. disagreeing without an explanation or clarification; irrelevant response

Memory

1 point - mention of any communication provider e.g. BT, broadband, or line rental company.

0 points – don't know, can't remember or incorrect recall.

Persuasion

2 points - reference to her desires, beliefs or intention to affect his actions or feelings e.g. she wants him to stop smoking, she wants him to not smoke around her, she is trying to make him feel guilty.

1 point -facts (he is smoking, she is pregnant, he is going to be a dad, to stop), outcomes (to stop him smoking), statements about him being a father and smoking (its irresponsible, he needs to grow up, its not a good influence on the child, its bad for his health), but do not reference her intentions.

0 points - incorrect or irrelevant facts or factors.

2 point mental state answer :

She wants to make him feel bad for smoking

What Next?

2 point - statement that responds to Alice's wanting him to give up or showing a desire or commitment to quit even if smoking now.

1 point -response that shows will change behavior e.g. smoking outside, without acknowledgement of Alice wanting to stop. Minimal responses that are still socially appropriate e.g. you're right. Asking for Alice to sympathise/minimizing e.g. its only one, without acknowledgement of wanting to quit.

0 points -don't know, socially inappropriate e.g. its my decision or my choice, or irrelevant comments that don't respond to Alice's previous statement e.g. Ill try.

Memory

1 point - a father, dad.

0 points - don't know, can't remember or incorrect recall.

Joke

Why?

2 points – he is joking or an explanation of the joke e.g. he thinks the politicians are full of rubbish or are rubbish.

1 point – facts (he is commenting on politics or situation), feelings (he is annoyed), traits (he is being cynical), Max's intentions/opinion of politicians that miss desire to be humorous (politicians lie, he doesn't like politicians, he thinks the politicians should change, politicians don't tell the truth, discredit the party).

2 point mental state answer:

He wants Alice to know that he thinks the politicians are full of rubbish

What Next?

2 point – statement that acknowledges joke and either agrees, asks for clarification, more information or challenges his opinion stated in the joke; responds to his joke with a second appropriate joke e.g. 'I hate talking politics with you'.

1 point – simple description of what would say or do e.g. laugh, sigh, I agree with you; statement that makes no explicit reference to the presence of a joke.

0 points – don't know, response that shows no understanding of the joke e.g. what does that mean? or directly challenges him missing the point of the joke e.g. 'that's not the answer to the question' or irrelevant comments.

Memory

1 point – newspaper, paper or reading.

0 points – don't know, can't remember or incorrect recall.

Double Bluff

2 point – mention of double bluff or reference to Max saying the truth in such a way (e.g. sarcasm) that Alice will think he is joking/not telling the truth or an expression that conveys this e.g. hiding in plain sight.

1 point -reference to him joking/trying to be funny without reference to Alice not believing him, facts (he was looking at a dating website, it's what he was doing), feelings (guilty), or trying to mislead her e.g. pretend doing nothing or play an emotional game.

0 points – incorrect e.g. to try and break up with Alice, he is overly honest or is shocking her. Irrelevant answers.

2 point mental state answer:

Max wanted Alice to think he was joking so she wouldn't believe that he would be looking at an online dating website.

What Next?

2 point – statement that assumes he is joking and/or makes a second joke in response e.g. 'I have been looking for another boyfriend'.

1 point – simple description of what would say or do e.g. laugh. Socially appropriate response that doesn't acknowledge joke e.g. how was your day?'. Response that asks for clarification whether he was joking.

0 points –don't know, response that assumes statement is sincere, socially inappropriate to the situation or irrelevant comments.

Memory

1 point – dating website or name of website or response that implies other women.

0 points – don't know, can't remember or incorrect recall.

Pretence

2 points -reference to Alice pretending they were rowing a boat; they are playing make believe/imaginary game; role play.

1 point - facts (they bumped into each other), simple intention (she is joking, flirtatious, affectionate, being surreal), states (she is being silly, playful, amused), actions (messing around, playing a game) or just stating they are rowing without mental state words e.g. pretence.

0 points – incorrect or irrelevant response.

2 point mental state answer:

Alice was playing make believe and wanted Max to think he had splashed her.

What Next?

2 point – statement that continues the make believe scenario or makes a joke out of the scenario or appropriate pretence through action e.g. pretending to splash with an imaginary oar.

1 point – simple description of what would say or do that makes no acknowledgement of joke e.g. I really enjoyed that

0 points –Don't know, response that understands comment as the truth e.g. apologizes sincerely, socially inappropriate or irrelevant comments.

Memory

1 point – mention of a chair.

0 points – don't know, can't remember or incorrect recall.

White Lie -

2 points - reference to white lie or making her feel good or not wanting to hurt Alice's feelings

1 point - response that states simple traits (he is nice, being supportive, polite) or is simply relational (he likes her). Incomplete response (offering fake praise) or solely motivational (so she won't be annoyed, avoid an argument, reassure her).

0 points – incorrect e.g. 'he thought it was good' or only 'he didn't like it', or irrelevant responses.

2 point mental state answer:

He doesn't want to hurt her feelings

What Next?

2 points – statement that acknowledges that Max's comment might not have been completely honest *and* either asks for additional clarification or additional feedback in socially appropriate manner 'do you really mean that?'; sarcastic agreement with his opinion that implies it could be improved.

1 point – Incomplete response e.g. thank you, that doesn't appreciate white lie.

0 points – don't know, socially inappropriate e.g. response that sees comment as unsupportive or misses intention of white lie, or irrelevant comments.

Memory

1 point – mentions guitar.

0 points – don't know, can't remember or incorrect recall.

Control Plant

2 points -response that states the tropical plant requires sunlight, warmth and humidity (2 needed) and that in the bathroom the plant will get these due to it being south facing.

1 point - reference to only warmth, sunlight or humidity due to either it being south facing or the bathroom. Reference to facts (it's a tropical plant, she's giving advice, it'll look nice); simple intentions (she is being helpful, letting him know she agrees, supportive, she thinks it's the best place) incomplete answers that do not mention important factors (it will grow there).

0 points - incorrect e.g. she doesn't like the plant, or irrelevant responses.

What Next?

2 points - statement that shows agreement with option chosen or provides alternative that has a rational e.g. kitchen windows.

1 point -Simple alternative without explanation.

0 points - don't know, socially inappropriate e.g. disagreeing without an explanation, or irrelevant response.

2 point mental state answer:

She thinks the upstairs bathroom will be the best environment for the plant to survive in and for Max to know she is being helpful.

Memory –

1point - mentions the bathroom.

0 points – don't know, can't remember or incorrect recall

6.9 Information Sheets and Consent Forms for SSFt Phase 2.

Institute of Psychiatry
at the Maudsley

South London and Maudsley 
NHS Foundation Trust

KING'S
College
LONDON

University of London

Department of Psychology
Institute of Psychiatry
De Crespigny Park
London, SE5 8AF

INFORMATION SHEET – ASD GROUP

Strange Situations: A Video-Based Test of Social Understanding.

We would like to invite you to take part in a research study looking at the ability to understand other people's thoughts and feelings. Before you decide whether to take part in this study, please take time to read the following information carefully. Please ask if there is anything that is not clear to you or if you would like more information.

Why is the study being done?

The study is a student research project, being carried out as part of a Doctorate in Clinical Psychology at the Institute of Psychiatry.

People with Autism Spectrum Disorders (ASDs) often say that they find social situations more confusing than their friends and family who don't have an ASD. This research is looking into why this might be. We have created a series of short films that show situations that people are faced with everyday and might be confusing or strange. By telling us what you think of the films we will better understand how people with an ASD view the social world when compared to people of a similar age who do not have an ASD.

This study is important because at the moment there are not many tools that are specifically designed to look at social understanding in adulthood. By taking part, you will

help us to understand more about how adults with ASD perceive the social world, which could in turn helps us to improve assessments and interventions for adults with an ASD in the future.

Why have I been chosen?

You have been invited to participate in this study because you have a diagnosis of an ASD (e.g. autism, Asperger syndrome).

Do I have to take part?

It is completely up to you whether you take part in the study or not. If you decide to take part, you can change your mind at any time and leave the study without giving a reason and only the data collected up until that point will be used. Refusal to take part in this study will not in any way affect you or affect the standard of any care you receive. If you have any questions about this project, please ask the researcher before you decide whether to take part.

What do I have to do if I agree to take part?

If you decide to take part, we will ask you to complete some questionnaires before attending and sign a consent form when you arrive. Signing the consent form shows that you understand the study and are willing to take part. You will then be asked to view some short videos and answer questions about them.. You will also be asked to complete some tasks that look at how you understand social situations and other people's emotions. Some of the tasks will be completed on computer and others will involve using pencil and paper. The questionnaires and tasks are not a test; there are no right or wrong answers. The appointment will take about 90 minutes in total to complete. If you attend this appointment you will receive £10 for taking part and will be reimbursed your exact travel costs (however this will be negotiated should your travel costs dramatically exceed £5).

What are the possible benefits from the research?

There is no direct benefit to you from taking part in this study. However, we hope the information gathered from the study will help us understand more about ASD in adulthood.

Will my taking part in this study be kept confidential?

Any information you provide will be kept private and will not be shown to anybody apart from the researchers. Your data will be stored without your name on it, on a secure computer or in a locked cabinet at the Institute of Psychiatry. Regulatory bodies may wish to view this anonymised information for audit purposes only.

What will happen to the results of the research study?

The results of the study will be written up in a thesis as part of a Doctorate in Clinical Psychology. The results will also be submitted for publication in academic psychology journals. No personal information will be identified in any publication of the results. The final results of the study will be available to you on request. If the results are published the anonymised data will be stored for 4 years for validation purposes and then disposed of securely.

Who has reviewed the study?

This research has been reviewed and approved by the [insert info]

Contact for further information?

If you have any further questions please feel free to call or email me at any time.

Contact information:

Mr Kim Murray

Trainee Clinical Psychologist

Tel: 0207 848 0223

Email: kim.k.murray@kcl.ac.uk

Address: Department of Psychology, Institute of Psychiatry, Addiction Sciences Building
4 Windsor Walk, London, SE5 8AF

Institute of Psychiatry
at the Maudsley

South London and Maudsley **NHS**
NHS Foundation Trust

KING'S
College
LONDON

University of London

Department of Psychology
Institute of Psychiatry
De Crespigny Park
London, SE5 8AF

INFORMATION SHEET – Healthy Control Group

Strange Situations: A Video-Based Test of Social Understanding.

We would like to invite you to take part in a research study looking at the ability to understand other people's thoughts and feelings. Before you decide whether to take part in this study, please take time to read the following information carefully. Please ask if there is anything that is not clear to you or if you would like more information.

Why is the study being done?

The study is a student research project, being carried out as part of a Doctorate in Clinical Psychology at the Institute of Psychiatry.

People with Autism Spectrum Disorders (ASDs) often say that they can find social situations more confusing than their friends and family who don't have an ASD. This research is looking into why this might be. We have created a series of short films that show situations that people are faced with everyday and might be confusing or strange. By telling us what you think of the films we will better understand how people with an ASD view the social world when compared to people of a similar age who do not have an ASD.

This study is important because at the moment there are not many tools that are specifically designed to look at social understanding in adulthood. By taking part, you will help us to understand more about how adults with ASD perceive the social world, which could in turn help us to improve assessments and interventions for adults with an ASD in the future.

Why have I been chosen?

You have been invited to participate in this study because you are an adult who does not have a diagnosis of an ASD.

Do I have to take part?

It is completely up to you whether you take part in the study or not. If you decide to take part, you can change your mind at any time and leave the study without giving a reason and only the data collected up until that point will be used. Refusal to take part in this study will not in any way affect you or affect the standard of any care you receive. If you have any questions about this project, please ask the researcher before you decide whether to take part.

What do I have to do if I agree to take part?

If you decide to take part, we will ask you to complete a questionnaire. You may be asked to come for another appointment at the Institute of Psychiatry or this may be all we require from you. If you come for an appointment you will have to complete some further questionnaires before attending and sign a consent form when you arrive. Signing the consent form shows that you understand the study and are willing to take part. You will then be asked to view some short videos and answer questions about them. You will also be asked to complete some tasks that look at how you understand social situations and other people's emotions. Some of the tasks will be completed on computer and others will involve using pencil and paper. The questionnaires and tasks are not a test; there are no right or wrong answers. The appointment will take about 90 minutes in total to complete. If you attend this appointment you will receive £10 for taking part and will be reimbursed your exact travel costs (however this will be negotiated should your travel costs dramatically exceed £5)

What are the possible benefits from the research?

There is no direct benefit to you from taking part in this study. However, we hope the information gathered from the study will help us understand more about ASD in adulthood.

Will my taking part in this study be kept confidential?

Any information you provide will be kept private and will not be shown to anybody apart from the researchers. Your data will be stored without your name on it, on a secure computer or in a locked cabinet at the Institute of Psychiatry. Regulatory bodies may wish to view this anonymised information for audit purposes only.

What will Happén to the results of the research study?

The results of the study will be written up in a thesis as part of a Doctorate in Clinical Psychology. The results will also be submitted for publication in academic psychology journals. No personal information will be identified in any publication of the results. The final results of the study will be available to you on request. If the results are published the anonymised data will be stored for 4 years for validation purposes and then disposed of securely.

Who has reviewed the study?

This research has been reviewed and approved by the National Research Ethics Service Committee London – Westminster (13/L0/0092) and the Institute of Psychiatry's/South London and Maudsley's Research and Development Office (R&D2013/016).

Contact for further information?

If you have any further questions please feel free to call or email me at any time.

Contact information:

Mr Kim Murray

Trainee Clinical Psychologist

Tel: 07999 929 241

Email: kim.k.murray@kcl.ac.uk

Address: Department of Psychology, Institute of Psychiatry, Addiction Sciences Building
4 Windsor Walk, London, SE5 8AF

Institute of Psychiatry
at the Maudsley

South London and Maudsley **NHS**
Foundation Trust

KING'S
College
LONDON

University of London

Department of Psychology
Institute of Psychiatry
De Crespigny Park
London, SE5 8AF

CONSENT FORM FOR PARTICIPANT

Strange Situations: A Video-Based Test of Social Understanding.

If you wish to take part in the above study, please read and complete the section below.

This is to confirm that I freely agree to take part in the above research. The researcher has explained to me why the study is taking place, what I will be asked to do and how long this will take. I have read the information section and understand the nature of the study. I understand that the study is purely a research project, and I do not expect to gain any personal benefit from taking part.

I am aware that the study involves completing some questionnaires, as well as taking part in pencil and paper tasks and watching short clips on the computer and then telling the researcher about what I understood about the short clips.

I understand that I am free to leave the research study at any time without giving a reason. I also understand that if I refuse to take part in the above study, this will not in any way affect any care I receive. I understand that the information I give is strictly confidential and will not be made publicly available. All information is for research purposes and will not be identifiable as mine.

Please initial boxes

1. I confirm that I have read and understand this information and consent form and have had the opportunity to ask questions.

☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

☐

I agree to be contacted for both future research related to this study and other pieces of research.

☐

I agree to take part in the above study.

☐

5. I agree to be audio recorded.

☐

Name of Participant

Date

Signature

I confirm that the project has been explained to the participant:

Name of Researcher

Date

Signature

**National Research Ethics
Service**



NRES Committee London - Westminster
Research Health Authority
Ground Floor, Skipton House
80 London Road
SE1 6LH

Telephone:
Facsimile:

12 February 2013

Mr Kim Murray
Trainee Clinical Psychologist
Camden and Islington NHS Foundation Trust
Institute of Psychiatry, King's College London
3rd Floor, Addiction Sciences Building,
4 Windsor Walk, London
SE5 8AF

Dear Mr Murray

Study title: Strange Situations: A Video Based Test of Social Understanding.
REC reference: 13/LQ/0092
Protocol number: N/A
IRAS project ID: 121593

Thank you for your recent letter, e-mailed on 11th February 2013. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 04 February 2013

Documents received

The documents received were as follows:

| Document | Version | Date |
|--|---------|------------------|
| Covering Letter | undated | |
| Participant Consent Form | 2 | 08 February 2013 |
| Participant Information Sheet: ASD Group | 2 | 08 February 2013 |

Approved documents

The final list of approved documentation for the study is therefore as follows:

| Document | Version | Date |
|------------------------------------|---------|------------------|
| Covering Letter | 1 | 19 December 2012 |
| Evidence of Insurance or Indemnity | | |

| | | |
|--|----------------------------------|------------------|
| Letter of Invitation to participant | 1 | 03 December 2012 |
| Other: CAG approval | | |
| Other: CV second supervisor | | |
| Participant Consent Form: consent to be contacted | 1 | 03 December 2012 |
| Participant Consent Form | 2 | 08 February 2013 |
| Participant Information Sheet: ASD Group | 2 | 08 February 2013 |
| Participant Information Sheet: Healthy Control Group | 2 | 08 February 2013 |
| Protocol | 1 | 03 December 2012 |
| REC application | 1 | 19 December 2012 |
| Referees or other scientific critique report | | |
| Summary/Synopsis | Patients and healthy controls V1 | 03 December 2012 |

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

13/LO/0092

Please quote this number on all correspondence

Yours sincerely

Alison O'Kane

Mrs. Alison O'Kane
Committee Co-ordinator

E-mail: NRESCCommittee.London-Westminster@nhs.net

Copy to: *Jenny Liebscher, KINGS COLLEGE LONDON*

6.10 The TAS-20

Sex: M / F

Age:

Date:

ID #:

T A S – 20

Using the scale provided as a guide, indicate how much you agree or disagree with each of the following statements by circling the corresponding number. Give only one answer for each statement.

Circle 1 if you **STRONGLY DISAGREE**
 Circle 2 if you **MODERATELY DISAGREE**
 Circle 3 if you **NEITHER DISAGREE NOR AGREE**
 Circle 4 if you **MODERATELY AGREE**
 Circle 5 if you **STRONGLY AGREE**

| | Strongly Disagree | Moderately Disagree | Neither Disagree Nor Agree | Moderately Agree | Strongly Agree |
|---|----------------------|------------------------|----------------------------------|---------------------|-------------------|
| 1. I am often confused about what emotion I am feeling. | 1 | 2 | 3 | 4 | 5 |
| 2. It is difficult for me to find the right words for my feelings. | 1 | 2 | 3 | 4 | 5 |
| 3. I have physical sensations that even doctors don't understand. | 1 | 2 | 3 | 4 | 5 |
| 4. I am able to describe my feelings easily. | 1 | 2 | 3 | 4 | 5 |
| 5. I prefer to analyze problems rather than just describe them. | 1 | 2 | 3 | 4 | 5 |
| 6. When I am upset, I don't know if I am sad, frightened, or angry. | 1 | 2 | 3 | 4 | 5 |
| 7. I am often puzzled by sensations in my body. | 1 | 2 | 3 | 4 | 5 |
| 8. I prefer to just let things happen rather than to understand why they turned out that way. | 1 | 2 | 3 | 4 | 5 |
| 9. I have feelings that I can't quite identify. | 1 | 2 | 3 | 4 | 5 |
| 10. Being in touch with emotions is essential. | 1 | 2 | 3 | 4 | 5 |

Date:

ID #:

T A S – 20

| | Strongly Disagree | Moderately Disagree | Neither Disagree Nor Agree | Moderately Agree | Strongly Agree |
|---|----------------------|------------------------|----------------------------------|---------------------|-------------------|
| 11. I find it hard to describe how I feel about people. | 1 | 2 | 3 | 4 | 5 |
| 12. People tell me to describe my feelings more. | 1 | 2 | 3 | 4 | 5 |
| 13. I don't know what's going on inside me. | 1 | 2 | 3 | 4 | 5 |
| 14. I often don't know why I am angry. | 1 | 2 | 3 | 4 | 5 |
| 15. I prefer talking to people about their daily activities rather than their feelings. | 1 | 2 | 3 | 4 | 5 |
| 16. I prefer to watch "light" entertainment shows rather than psychological dramas | 1 | 2 | 3 | 4 | 5 |
| 17. It is difficult for me to reveal my innermost feelings, even to close friends. | 1 | 2 | 3 | 4 | 5 |
| 18. I can feel close to someone, even in moments of silence. | 1 | 2 | 3 | 4 | 5 |
| 19. I find examination of my feelings useful in solving personal problems. | 1 | 2 | 3 | 4 | 5 |
| 20. Looking for hidden meanings in movies or plays distracts from their enjoyment. | 1 | 2 | 3 | 4 | 5 |

6.11 The IRI

INTERPERSONAL REACTIVITY INDEX

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E. When you have decided on your answer, fill in the letter on the answer sheet next to the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

ANSWER SCALE:

| | | | | |
|-------------|---|---|---|--------------|
| A | B | C | D | E |
| DOES NOT | | | | DESCRIBES ME |
| DESCRIBE ME | | | | VERY |
| WELL | | | | WELL |

1. I daydream and fantasize, with some regularity, about things that might Happén to me.
2. I often have tender, concerned feelings for people less fortunate than me.
3. I sometimes find it difficult to see things from the "other guy's" point of view.
4. Sometimes I don't feel very sorry for other people when they are having problems.
5. I really get involved with the feelings of the characters in a novel.
6. In emergency situations, I feel apprehensive and ill-at-ease.

I am usually objective when I watch a movie or play, and I don't often get completely caught up in it.

8. I try to look at everybody's side of a disagreement before I make a decision.

9. When I see someone being taken advantage of, I feel kind of protective towards them.

10. I sometimes feel helpless when I am in the middle of a very emotional situation.

I sometimes try to understand my friends better by imagining how things look from their perspective.

12. Becoming extremely involved in a good book or movie is somewhat rare for me.

13. When I see someone get hurt, I tend to remain calm.

14. Other people's misfortunes do not usually disturb me a great deal.

If I'm sure I'm right about something, I don't waste much time listening to other people's arguments.

16. After seeing a play or movie, I have felt as though I were one of the characters.

17. Being in a tense emotional situation scares me.

When I see someone being treated unfairly, I sometimes don't feel very much pity for them.

19. I am usually pretty effective in dealing with emergencies.

20. I am often quite touched by things that I see Happén.

21. I believe that there are two sides to every question and try to look at them both.

22. I would describe myself as a pretty soft-hearted person.

23. When I watch a good movie, I can very easily put myself in the place of a leading character.

24. I tend to lose control during emergencies.

25. When I'm upset at someone, I usually try to "put myself in his shoes" for a while.

When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were Happening to me.

27. When I see someone who badly needs help in an emergency, I go to pieces.

28. Before criticizing somebody, I try to imagine how I would feel if I were in their place.

NOTE: (-) denotes item to be scored in reverse fashion

PT = perspective-taking scale

FS = fantasy scale

EC = empathic concern scale

PD = personal distress scale

A = 0

B = 1

C = 2

D = 3

E = 4

Except for reversed-scored items, which are scored:

A = 4

B = 3

C = 2

D = 1

E = 0

6.12 Strange Stories and Scoring

Strange Stories (Happé 1994)

Late one night old Mrs. Peabody is walking home. She doesn't like walking home alone in the dark

because she is always afraid that someone will attack her and rob her. She really is a very nervous

person! Suddenly, out of the shadows comes a man. He wants to ask Mrs. Peabody what time it is, so he

walks toward her. When Mrs. Peabody sees the man coming toward her, she starts to tremble and says,

“Take my purse, just don't hurt me please!”

Q: Why did she say that?

Simon is a big liar. Simon's brother Jim knows this, he knows that Simon never tells the truth!

Now yesterday Simon stole Jim's ping-pong paddle, and Jim knows Simon has hidden it somewhere,

though he can't find it. He's very cross. So he finds Simon and he says, “Where is my pingpong

paddle? You must have hidden it either in the cupboard or under your bed, because I've

looked everywhere else. Where is it, in the cupboard or under your bed?” Simon tells him the

paddle is under his bed.

Why will Jim look in the cupboard for the paddle?

Brian is always hungry. Today at school it is his favourite meal—sausages and beans. He is a very

greedy boy, and he would like to have more sausages than anybody else, even though his mother

will have made him a lovely meal when he gets home! But everyone is allowed two sausages and

no more. When it is Brian's turn to be served, he says, "Oh, please can I have four sausages, because

I won't be having any dinner when I get home!"

Q: Why does Brian say this?

One day Aunt Jane came to visit Peter. Now Peter loves his aunt very much, but today she is

wearing a new hat; a new hat which Peter thinks is very ugly indeed. Peter thinks his aunt looks silly

in it, and much nicer in her old hat. But when Aunt Jane asks Peter, "How do you like my new hat?,"

Peter says, "Oh, its very nice."

Q: Why does he say that?

During the war, the Red army captures a member of the Blue army. They want him to tell them where his army's tanks are; they know they are either by the sea or in the mountains. They know that the

prisoner will not want to tell them, he will want to save his army, and so he will certainly lie to them.

The prisoner is very brave and very clever, he will not let them find his tanks. The tanks are really in

the mountains. Now when the other side asks him where his tanks are, he says, "They are in the

mountains."

Q: Why did the prisoner say that?

Helen waited all year for Christmas, because she knew at Christmas she could ask her parents for a

rabbit. Helen wanted a rabbit more than anything in the world. At last Christmas Day arrived, and

Helen ran to unwrap the big box her parents had given her. She felt sure it would contain a little rabbit

in a cage. But when she opened it, with all the family standing round, she found her present was

just a boring old set of encyclopedias, which Helen did not want at all! Still, when Helen's parents

asked her how she liked her Christmas present, she said, "It's lovely, thank you. It's just what I

wanted."

Q: Why did she say this?

Jill wanted to buy a kitten, so she went to see Mrs. Smith, who had lots of kittens she didn't want.

Now Mrs. Smith loved the kittens, and she wouldn't do anything to harm them, though she couldn't keep them all herself. When Jill visited she wasn't sure she wanted one of Mrs. Smith's

kittens, since they were all males and she had wanted a female. But Mrs. Smith said, "If no one

buys the kittens I'll just have to drown them!"

Q: Why did Mrs. Smith say that?

A burglar who has just robbed a shop is making his getaway. As he is running home, a policeman

on his beat sees him drop his glove. He doesn't know the man is a burglar, he just wants to tell him

he dropped his glove. But when the policeman shouts out to the burglar, "Hey, you! Stop!," the

burglar turns round, sees the policeman and gives himself up. He puts his hands up and admits that

he did the break-in at the local shop.

Q: Why did the burglar do that?

Strange Stories Scoring from White et al., (2009)

Misunderstanding (Peabody)

2 points—reference to her belief that he was going to mug her or her ignorance of his real intention

1 point—reference to her trait (she's nervous) or state (she's scared) or intention (so he wouldn't hurt her) without suggestion that fear was unnecessary

0 points—factually incorrect/irrelevant answers; reference to the man actually intending to attack her

Double Bluff (Ping pong paddle)

2 points—reference to Jim knowing Simon lies

1 point—reference to facts (that's where it really is, Simon's a big liar) or Simon hiding it without reference to implications of lying

0 points—reference to general nonspecific information (because he looked everywhere else)

Persuasion (Sausages)

2 points—reference to fact that he's trying to elicit sympathy, being deceptive

1 point—reference to his state (greedy), outcome (to get more sausages) or factual

0 points—reference to a motivation that misses the point of sympathy elicitation / deception, or factually incorrect

White Lie (Hat)

2 points—reference to white lie or wanting to spare her feelings; some implication that this is for aunt's benefit rather than just for his, desire to avoid rudeness or insult

1 point—reference to trait (he's a nice boy) or relationship (he likes his aunt); purely motivational (so she won't shout at him) with no reference to aunt's thoughts or feelings; incomplete explanation (he's lying, he's pretending).

0 points—reference to irrelevant or incorrect facts/feelings (he likes the hat, he wants to trick her)

Double Bluff (army)

2 points—reference to 1 that other army will not believe and hence look in other place, reference to prisoner's realization that that's what they'll do, or reference to double bluff

1 point—reference to outcome (to save his army's tanks) or to mislead them

0 points—reference to motivation that misses the point of double bluff (he was scared)

White Lie (RABBITT)

2 points—reference to white lie or wanting to spare their feelings; some implication that this is for parent's benefit rather than just for her, desire to avoid rudeness or insult

1 point—reference to trait (she's a nice girl) or relationship (she likes her parents); purely motivational (so they won't shout at her) with no reference to parent's thoughts or feelings; incomplete explanation (she's lying, she's pretending)

0 points—reference to irrelevant or incorrect facts/feelings (she likes the present, she wants to trick them)

Persuasion (Kittens)

2 points—reference to persuasion, manipulating feelings, trying to induce guilt / pity

1 point—reference to outcome (to sell them or get rid of them in a way which implies not drowning) or simple motivation (to make Jill sad)

0 points—reference to general knowledge or dilemma without realization that the statement was not true (she's a horrible woman)

Misunderstanding (Burglar)

2 points—reference to belief that policeman knew that he'd burgled the shop

1 point—reference to something factually correct in story

0 points—factually incorrect / irrelevant answers

6.13 SSFt instructions, Scoring Sheet and Order A (Order B is reversed)

You are about to watch a series of short video clips. In each clip there will usually be the same two characters; a man called Max and a woman called Alice. Max and Alice are boyfriend and girlfriend, and live and work together. Each video clip is self-contained, and what happens in one clip doesn't relate to what happens in other clips; there isn't a story running from one clip to another.

I will show you each clip only once and after each clip has finished I will ask you three questions about what you have just seen. I will write down your answers and also audio record them in case I miss anything. Some of the questions do not have right or wrong answers – I'm interested in what you think is going on in the clips, so please just answer the questions as best you can..

Do you have any questions?

[If participant asks questions regarding the content of the material e.g. "How long have they been in a relationship?" Respond with

Please give your answers based just on what you see and the information I gave you at the beginning. Would you like me to re-read the instructions?

OK. The first three clips you are going to watch are just a practice so that you can see what is involved. We will watch the first clip twice. Please place the headphones in your ears and tell me if you need the volume adjusting [play practice control clip and adjust volume accordingly]. Ok, now I will ask you three questions. afterwards. (Play practice control clip again)

NB : No prompting to be provided throughout unless participant's response is unintelligible or they use words that the research does not understand. In such cases ask :

'I am sorry but I didn't hear/understand your answer, could you please repeat it for me?'

Practice 1 Forget (present)

Why did Alice say that?

If you were in Max's situation, what would you say next?

Where had Alice just come from?

Ok great, and here is the second practice clip – this time you'll see it only once, just like in the real thing [show practice experimental clip]

Practice 2 Control (cactus)

Why did Max say that?

If you were in Alice's situation, what would you say next??

How long was Max going away for?

Ok great, and here is the second practice clip – this time you'll see it only once, just like in the real thing [show practice experimental clip]

Practice 3 Persuasion (mum)

Why did Max say that?

If you were in Alice's situation, what would you say next??

Who does Alice find really difficult?

Do you have any questions?

Now you are going to watch a series of clips similar to the ones you have just seen and I will ask you to answer some questions after you have viewed each clip just once.

Play all clips and following each pause the task and ask

Item 1 Idiom (tango)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Max say John should do?

Item 2 Control (Soup)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max going to cook the next day?

Item 3 Mixed Emotions (biking accident)

Why did Alice say everything she said in the clip?

If you were in Max's situation, what would you say next??

What was the disgusting picture of?

Item 4 Misunderstanding (burglar)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Max doing when Alice called out?

Item 5 Lie (Pub)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Where was Alice when she spoke to Max on the phone?

Item 6 Forget (potatoes)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Alice buy from the shop?

Item 7 AR (dress)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Alice about to put on Max?'

Item 8 Irony (injection)

Why did Max say that?]

If you were in Alice's situation, what would you say next??

What was Max doing to his tummy?.

Item 9 Control (Line Rental)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Who was Max talking to on the phone?

Item 10 Persuasion (smoking)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What did Alice say Max was going to be?

Item 11 Joke (politics)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max reading?

Item 12 DB (dating website)

Why did Max say that?

If you were in Alice's situation, what would you say next??

What was Max looking at on the computer at the start?

Item 13 Pretend (rowing)

Why did Alice say that?

If you were in Max's situation, what would you say next??

What was Alice sitting on?

Item 14 White Lie (song)

Why did Max say that?

If you were in Alice's situation, what would you say next??

'What instrument was Alice playing?

Item 15 Control (plant)

Why did Alice say that?

If you were in Max's situation, what would you say next??

Where did Alice say the plant could go?

6.14 Associations Between Traditional ToM Measures

Associations of both groups' performance on the traditional social cognition measures: Spearman's partial coefficients and (r^2 values).

| | SS PST | F-HT Ac | F-HT PST | Eyes | TASIT |
|--------------------------|--------------------|---------|---------------------|------|-------|
| <i>ASD [N = 20]</i> | | | | | |
| SS Ac | .51* (.26) | .08 | .21 | .16 | .36 |
| SS PST | | .15 | .17 | .15 | .06 |
| F-HT Ac | | | .78**a (.61) | -.40 | .02 |
| F-HT PST | | | | -.24 | -.08 |
| Eyes | | | | | .44 |
| <i>Controls [N = 20]</i> | | | | | |
| SS Ac | .87** (.76) | .32 | .14 | .27 | .08 |
| SS PST | | .38 | .30 | .30 | .22 |
| F-HT Ac | | | .74**a (.55) | .34 | .24 |
| F-HT PST | | | | .26 | .12 |
| Eyes | | | | | .37 |

^a Pearson's's partial r . * = p -value < .05. ** = p -value < .01. Ac= Accuracy.

6.15 Traditional ToM Measures and IRI, TAS-20, ADI-R and AQ Associations.

Association of performance across social cognition measures with alexithymia and empathy: Spearman's partial correlation coefficients controlling for VIQ.

| Traditional Social Cognition Measures | TAS-20 | | IRI | | |
|---------------------------------------|--------|------|-------|------|------|
| | Total | PT | F | EC | PD |
| ASD [<i>N</i> = 19] | | | | | |
| SS Accuracy | -.20 | .18 | -.49* | -.02 | .16 |
| F-HT Accuracy ^a | .08 | -.15 | -.18 | -.30 | -.04 |
| Eyes | .02 | -.18 | .04 | .01 | -.05 |
| TASIT | -.32 | -.18 | -.17 | -.09 | -.05 |
| Controls [<i>N</i> = 20] | | | | | |
| Traditional Social Cognition Measures | | | | | |
| SS Accuracy | -.03 | -.07 | .09 | .21 | .11 |
| F-HT Accuracy ^a | -.15 | .14 | -.01 | -.23 | .03 |
| Eyes | -.34 | .21 | .25 | -.05 | -.16 |
| TASIT | -.14 | .22 | .13 | .07 | -.34 |

a

. Pearson's *r*

Association of performance between social cognition measures and ASD symptoms:
partial Spearman's correlation coefficient controlling for VIQ.

| Traditional Social Cognition Measures | ASD Trait Measures | | |
|---------------------------------------|--------------------|---------------------|--------------|
| | AQ | ADI-R Communication | ADI-R Social |
| <i>ASD [N = 20]</i> | | | |
| SS Accuracy | .02 | -.29 | -.16 |
| F-HT Accuracy | .12 ^a | .05 | -.27 |
| Eyes | -.05 | .04 | .43 |
| TASIT | .06 | .20 | .41 |
| <i>Controls [N = 20]</i> | | | |
| SS Accuracy | -.41 | - | - |
| F-HT Accuracy | -.07 ^a | - | - |
| Eyes | -.24 | | |
| TASIT | -.43 | | |

^aPearson's r.

6.16 SSFt Order Effects

| The Strange Situations Film test (SSFt).... | Order A N=20 | Order B N=20 | <i>t</i> | <i>df</i> | <i>p-value</i> | <i>d</i> | 95% <i>CI</i> |
|---|-----------------|-----------------|----------|-----------|----------------|----------|---------------------------|
| Experimental Clips | | | | | | | |
| Intention (<i>max=24</i>) | 17.40 (3.64) | 17.00 (3.26) | 0.37 | 38 | .72 | 0.12 | -1.81 – 2.61 |
| Psychological State talk (<i>max=24</i>) | 12.10 (2.77) | 13.20 (2.46) | -1.33 | 38 | .192 | 0.41 | -2.68 – 0.46 ^a |
| Interaction (<i>max=24</i>) | 13.00 (5.19) | 14.85 (4.63) | -1.19 | 38 | .241 | 0.37 | 3.92 – 8.78 |
| Memory (<i>max=12</i>) | 11.55 (0.69) | 11.85 (0.37) | -1.73 | 29 | .095 | 0.55 | -0.66 – 0.26 ^a |
| Control Clips [<i>max = 24</i>] | | | | | | | |
| Intention (<i>max=6</i>) | 4.25 (1.41) | 4.25 (1.01) | 0.00 | 38 | 1.00 | 0.00 | -0.44 – 0.75 ^a |
| Psychological State talk (<i>max=6</i>) | 1.10 (1.17) | 1.10 (0.91) | 0.00 | 38 | 1.00 | 0.00 | -0.65 – 0.66 ^a |
| Interaction (<i>max=6</i>) | 5.25 (0.97) | 5.85 (0.49) | -2.48 | 28 | .02 | 0.78 | -1.08 – -0.14 |
| Memory (<i>max = 3</i>) | 3.00 (0.00) | 2.95 (0.22) | 1.00 | 19 | .32 | 0.32 | 0.04 – 0.18 ^a |

Within group differences on the interaction question yielded non-significant differences that were supported by an independent bootstrap analysis for the ASD group ($t(15) = -1.57, p = .138, CI = -1.33 - 1.6$) and the control group ($t(10) = -1.89, p = .08, CI = -1.22 - -.04$).

Differences in order A's increased 0 point responses across different clips was evenly spread: Clip 2 (+2) Clip 9 (+1) and Clip 15 (+1). This suggests that it was not a particular clips order that affected performance per se but individuals within group A as they were not randomized.

**Referral Rates, Clinical Presentations and
Outcomes of Cognitive Behaviour Therapy for
Obsessive-Compulsive Disorder in Young People
with and Without Autism Spectrum Disorders**

Mr. Kim Murray

Supervised by Dr Georgina Krebs

Service Evaluation Project

**Department of Psychology, Institute of Psychiatry
King's College London**

Abstract

Background: Autism Spectrum Disorders (ASDs) are a continuum of neurodevelopmental disorders and the prevalence of Obsessive Compulsive Disorder (OCD) is high within this population compared to typically developing young people.

Method: Retrospective referrals to a specialist pediatric OCD service at the Maudsley Hospital were examined. Prevalence rates of referrals of young people with OCD plus ASD over a four year time period were analysed. Referrals of young people with OCD plus ASD (ASD+OCD) were compared to those without ASD (NoASD+OCD) at assessment on a range of clinical measures and demographic variables. Finally, clinical outcomes of young people with ASD+OCD following a course of cognitive behavior therapy (CBT) were compared to an NoASD+OCD group matched on age, gender and pre-treatment OCD symptom severity.

Results: The results suggest that referral rates have remained relatively stable. Young people with ASD tended to have more severe OCD symptoms as rated by clinicians and parent informants relative to those without ASD. Moreover, the ASD+OCD group responded less well to Cognitive Behavioural Therapy delivered within the clinic's protocol as compared to the matched NoASD+OCD group.

Conclusions: Overall, the findings support the importance of multiple informants at assessment and that CBT is beneficial for this client group but modified protocols are needed within the clinic when treating young people with these two conditions.

Table of Contents

| | |
|---|------------|
| List of Tables..... | 215 |
| List of Figures | 215 |
| 1 Introduction [Brief Literature Review] | 216 |
| 1.1 Definition and Prevalence of Autism Spectrum Disorders | 216 |
| 1.2 Anxiety Disorders in ASDs..... | 216 |
| 1.3 OCD in ASD..... | 217 |
| 1.4 Considerations When Delivering CBT for Young People with an ASD | 219 |
| 1.5 Treating Anxiety Disorders in ASD | 220 |
| 1.6 Treating OCD in ASDs | 221 |
| 1.7 Service Evaluation Aims and Hypotheses | 222 |
| 1.8 Service Context | 223 |
| 1.8.1 The National Specialist OCD and Related Disorders Service..... | 223 |
| 1.8.2 Assessment Process..... | 224 |
| 1.8.3 Cognitive Behavioural Therapy Protocol..... | 224 |
| 2 Method..... | 225 |
| 2.1 Participants..... | 225 |
| 2.2 Measures..... | 225 |
| 2.2.1 The Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS)..... | 225 |
| The Child Obsessive Compulsive Inventory-Revised (ChOCI-R)..... | 225 |
| 2.2.2 The Childrens' Global Assessment Scale (CGAS) | 226 |
| 2.3 Information Obtained..... | 226 |
| 2.4 Procedure..... | 226 |
| 3 Results | 227 |
| 3.1 Referral rates | 227 |
| 3.2 Demographics and Clinical Characteristics at Assessment | 228 |
| 3.3 Treatment..... | 231 |
| 4 Discussion..... | 234 |
| 4.1 Referral Rates..... | 234 |
| 4.2 Demographics and Initial Presentation | 235 |
| 4.3 Treatment Outcomes..... | 236 |
| 4.4 Treatment Modifications | 238 |
| 4.5 Limitations..... | 238 |
| 5 Service Recommendations | 241 |

| | |
|--|------------|
| 6 Dissemination of Results..... | 242 |
| 7 Reflections..... | 243 |
| 8 References..... | 245 |

List of Tables

| | |
|--|-----|
| Table 1: Matched Group's Initial Presentation..... | 231 |
|--|-----|

List of Figures

| | |
|--|-----|
| Figure 1: Frequency of referrals by group between 2007 and 2011..... | 227 |
| Figure 2: Relative percentage of each group when compared to the overall number of referrals per year between 2007 and 2011..... | 227 |
| Figure 3: Age of both groups and age of OCD onset in years..... | 229 |
| Figure 4: Percentage of each group meeting NSCT funding criteria and the percentage of males in each group..... | 229 |
| Figure 5: Total CY-BOCS, Child and Parent ChOCI-R, and CGAS scores at assessment..... | 230 |
| Figure 6: Mean total CY-BOCS score at the Beginning and End of treatment..... | 232 |
| Figure 7: Percentage of treatment responders (Responders) and remission rates as measured by CY-BOCS score of 14 or less (Remission CY-BOCs). | 233 |

1 Introduction [Brief Literature Review]

1.1 Definition and Prevalence of Autism Spectrum Disorders

Autism Spectrum Disorders (ASDs) are a continuum of neurodevelopmental disorders categorized by difficulties in reciprocal social interaction, communication, imagination and having restrictive/stereotyped repetitive interests (World Health Organization [WHO] International Classification of Diseases-10 [ICD-10]; 2008). Childhood Autism, Asperger's syndrome and Pervasive Developmental Disorders- Not Otherwise Specified (PDD-NOS) all fall under the umbrella term of ASDs and are differentiated by differences in symptomatology severity within the aforementioned triad, language delay and IQ (WHO, 2008). Traditionally, ASDs were believed to account for 0.05% of the population, however, more recent prevalence studies have suggested ASDs to be within 0.3-0.9% (Baird et al., 2006; Yeargin-Allsopp et al., 2003). Within the South Thames, London region, Baird, et al., (2006) identified prevalence rates of ASDs in 9-10 year olds to account for approximately 1% of the population, suggesting that ASDs are much more common than once thought. The same study found similar ratios of sex differences to US data (Yeargin-Allsopp, et al., 2003) with a male to female ratio of 3.3-1.

Whether these increased rates reflect the broadening of diagnostic categories or improved prevalence research, for mental health clinicians the issue of paramount importance is how to best assess and treat mental health referrals in the context of an ASD (White, Oswald, Ollendick, & Scahill, 2009). This is particularly true for clinicians working in the NHS, where policies such as clinical governance and an increasing drive towards 'Payment by Results' (Department of Health Payment by Results Team, 2012) mean that an increasing importance is being placed on clinical outcomes and ongoing improvement in service delivery. It is therefore imperative that services and clinicians working within them understand the clinical needs of young people with ASD and how best to treat distressing mental health problems associated with the condition.

1.2 Anxiety Disorders in ASDs

The most common forms of co-morbid Axis I disorders present in young people and adolescents with an ASD are the cluster of Anxiety Disorders (Ghaziuddin, 2002). Anxiety Disorders all share anxiety as their primary feature, but specific anxiety disorders manifest in distinct cognitive-behavioural reactions to feared stimuli e.g. avoidance of

potential contaminates in Obsessive Compulsive Disorder (OCD; Albano, Chorpita, & Barlow, 2003). Prevalence rates of anxiety disorders in typically developing children are estimated to range in community samples between 12-20% (Albano, et al., 2003). While research on a population level has not been conducted in youth with an ASD and co-morbid anxiety disorders, White, et al.,’s (2009) review suggests that 11-84% of young people are affected by anxiety that causes significant impairment. Simonoff et al., (2008) revealed that in a community sample cohort study of 112 10-14 year olds with ASD 41.9% of the sample met diagnosable criteria for at least one anxiety disorder with social anxiety being the most common (29.2%) and OCD occurring in 8.2% of cases. Similarly, de Bruin, Ferdinand, Meester, de Nijs & Verheij (2007) showed that in a clinic sample of 94, 6-12 year olds diagnosed with PDD-NOS, the single most common disorder diagnosed was simple phobia, with 6.2% of the sample meeting criteria for OCD.

1.3 OCD in ASD

OCD is an anxiety disorder characterized by intrusive, spontaneous and unwanted thoughts , images or urges (obsessions) that cause significant distress, which in turn lead to purposeful rituals; mental and/or behavioural (compulsions), that serve to reduce the distress experienced, but can in themselves become a significant contributor to the individual’s distress (Heyman, Mataix-Cols, & Fineberg, 2006). Symptoms have to cause the individual functional impairment and/or distress to warrant a diagnosis (WHO, 2008). Epidemiological studies, suggest rates of OCD in typically developing children to be between 1% of the population (Heyman, Mataix-Cols, & Fineberg, 2006). Thus, the prevalence rates of OCD of 6.2-8.2% found in ASD samples suggest that the disorder is overrepresented in ASD populations (de Bruin, et al., 2007; Simonoff, et al., 2008; White, et al., 2009).

OCD symptoms start most commonly at two distinct time points late childhood and early adulthood (Geller et al., 1998). Onset of co-morbid OCD in young people with an ASD has not been systematically evaluated to the author’s knowledge, but the phenomenological overlap of the two conditions, namely insistence on sameness and repetitive interests found in both ASD and OCD diagnoses, make differential diagnosis an important, but challenging consideration in the assessment of co-morbid OCD in young people with an ASD (Baron-Cohen, 1989). The presence of egodystonic self reports

regarding symptoms and the presence of obsessions if potential compulsive symptoms are present are therefore particularly important to consider (Wood & Gadow, 2010).

Insight into the excessive nature of compulsions and their egodystonic quality are hallmarks of adult OCD. While awareness criterion are more lenient when considering OCD in a typically developing paediatric population (Heyman, et al., 2006) when assessing OCD in paediatric ASD populations this calls in to question a number of factors. Firstly, alexythimic symptoms are more likely to be present in ASD populations (Tani et al., 2004) and rates of intellectual disability also, both of which can impact an individual's ability to describe internal experiences required for accurately assessing anxiety.

There has been controversy in the literature as to whether young people with ASD can therefore self report on anxiety symptoms reliably (Wood & Gadow, 2010). Mack et al., (2010) showed that when groups were matched on OCD severity, self and parent reports of OCD symptoms as measured by the Child Obsessive Compulsive Checklist (ChOCI; Uher, Heyman, Turner, & Shafran, 2008) were equivocal. However, the small N (12) in each group calls some of these findings into question as the authors note. In other studies, youth with ASD have shown to reliably self report, but also significantly underreport when compared to informant's (Shipman, Sheldrick, & Perrin, 2011). Furthermore, Gadow, Devincent, Sheldrick, & Perrin (2005), revealed that teachers can rate anxiety symptoms in their ASD pupils higher than parent-informants. Overall, the findings are mixed in this area.

Clinician reported data, as measured by the Child's Yale Brown Obsessive Compulsive Scale (CY-BOCS; Scahill, Riddle, McSwiggin-Hardin, & Ort, 1997) have shown that symptom profiles in youth with ASD and OCD compared with those with just an OCD diagnosis are broadly similar, but that there are some idiosyncrasies within each group (Lewin, Wood, Gunderson, Murphy, & Storch, 2011; Mack, et al., 2010). The data is inconsistent between studies however. For example reduced checking, repeating and washing compulsions within the ASD and OCD group were observed in one study (Lewin, et al., 2011) and not in another (Mack, et al., 2010). Of note, a trend towards reduced sexual obsessions in the ASD and OCD group was observed in both studies. In all, this evidence highlights both an interesting question as posed by Wood & Gadow (2010), 'Can self-report anxiety measures be deemed valid in paediatric ASD populations?', and secondly, the importance of multiple informants when assessing co-morbid anxiety and in

particular OCD in youth with ASD because of the phenomenological overlap between the two constructs.

1.4 Considerations When Delivering CBT for Young People with an ASD

Treating paediatric OCD has accumulated a significant evidence base in recent decades and the National Institute for Health and Care Excellence (NICE; 2005) recommend Cognitive Behavioural Therapy (CBT) that includes Exposure with Response Prevention (E/RP) as a first line treatment for young people with the condition. However, the presence of an ASD is typically an exclusion criterion in randomized controlled trials (RCTs) and the existing evidence and guidance relate to typically developing young people. There has been little research into the treatment of paediatric OCD in the context of ASD and there are no guidelines on what forms of therapy are best suited to this population.

Despite the limited research, there are a number of theoretical reasons to believe that young people with ASD may have difficulties engaging in standard, protocol-driven CBT interventions, which may require modification in order to optimize outcomes in this population. For example, difficulties in identifying thoughts, feelings and behaviours are a key component in CBT and ASD is associated with higher instances of alexithymia (Tani, et al., 2004) making this process potentially more challenging. The Theory of Mind hypothesis purports that young people with an ASD struggle to understand the thoughts, intentions and or beliefs of others (Baron-Cohen, Leslie, & Frith, 1985), which may affect the young person's ability to engage with discourse associated with CBT, e.g. 'what would your friend's make of that thought?' that require adopting another's perspective. Finally, executive functioning impairments associated with the condition (Hill, 2004) may impair session planning, implementation of homework and generating new and novel means to challenge fears, furthermore the literal and concrete linguistic style associated with the condition (Lord et al., 2000) may make the use of metaphors and hypothetical thought experiments more challenging, which are often used to educate young people about CBT principles.

1.5 Treating Anxiety Disorders in ASD

Over the last decade, there is a growing evidence base for the treatment efficacy of modified CBT in paediatric populations with both an ASD and an anxiety disorder in both group and individual formats (Lang, Regeister, Lauderdale, Ashbaugh, & Haring, 2010; White, et al., 2009).

In Lang et al.,’s (2010) systematic review assessing treatment of anxiety disorders in youth with an ASD they only identified one RCT (Wood et al., 2009), which could be considered ‘conclusive’ in its findings as the other three trials showed methodological flaws such as non-random assignment (Reaven et al., 2009) or the use of unblinded/therapists as raters (Chalfant, Rapee, & Carroll, 2007; Sofronoff, Attwood, & Hinton, 2005). All the other studies mentioned in the review did not include an experimental design (for review see Lang, et al., 2010). Since then, a number of trials have addressed these concerns (Reaven, Blakeley-Smith, Culhane-Shelburne, & Hepburn, 2012; Storch et al., 2013). In the well controlled trials, response rates to CBT were between 50-76% and remission rates ranged from 38-52.9% (Reaven, et al., 2012; Storch, et al., 2013; Wood, et al., 2009). While the approach taken was transdiagnostic, including treating separation-, social-, generalized- anxiety disorder, specific phobias and OCD there was no effect of initial diagnosis on treatment outcome and the above remission rates are similar to those found in CBT for paediatric OCD in neurotypical samples (Pediatric OCD Treatment Study [POTS] Team., 2004). However, the rates of OCD referrals were either low (Storch, et al., 2013) or not included (Reaven, et al., 2012) in the above trials.

Across all the trials CBT was the only therapeutic framework adopted, but protocols were modified to meet the needs of an ASD population, for example, the use of increased visual aids in treatment, greater parental involvement, using restricted interests to engage participants in treatment (Lang, et al., 2010) and in some cases addressing core autism symptomatology as part of the treatment package (Sze & Wood, 2008; Wood, et al., 2009). Parental involvement as a co-therapist or coach was a particularly important modification aimed at improving problems with generalizing skills learned inside the therapy room to the young person’s wider social environment, which was tested explicitly in one trial and shown to improve outcome over follow-up in group CBT (Reaven, et al., 2012). Parents were more heavily involved at all stages of trials investigating individual (family-based) CBT treatment also, for example during ERP, psychoeducation, social skills, encouraging

independence, implementing rewards and advocating within schools (Storch, et al., 2013; Wood, et al., 2009). Despite the clear theoretical rationale for involving parents, additional clinical improvements did not appear at three-month follow-up in a recent well controlled study of family-based CBT compared to treatment as usual (TAU) (Storch, et al., 2013).

The overall heterogeneity of clinical presentations, severity, treatment format and number of sessions makes comparative conclusions difficult. These points to a growing need for equivalence trials as so far the majority of trials include only a TAU control arm rather than an active treatment group. For example, at present it is unclear whether group or individual treatments are more effective or whether this population requires additional sessions compared to typically developing populations. To date, only one equivalence RCT looking at CBT specifically for OCD in adults and adolescents with an ASD has been conducted with the authors finding both CBT and anxiety management equivocal treatment packages (Russell et al., 2013).

1.6 Treating OCD in ASDs

The evidence for treating OCD specifically in an ASD paediatric population is lacking and is limited to just two case studies (Lehmkuhl, Storch, Bodfish, & Geffken, 2008; Reaven & Hepburn, 2003) and some participants included in the aforementioned RCT (Russell, et al., 2013). In Reaven & Hepburn's (2003) case, the young person's OCD severity reduced over 14 sessions from the moderate range to the normal range. The intervention was based on March and Mulle's (1998) evidenced based protocol, but the authors discuss numerous modifications necessary for working therapeutically with a young person with an ASD. Parents were more actively involved throughout the psychoeducation phase and abstract concepts were made more concrete, such as creating a cardboard cut-out model of a CBT tool box with cardboard tools within it. During hierarchy formation and subsequent ERP the young person's language was adopted and the use of token rewards was implemented to tackle reduced intrinsic social motivators for overcoming OCD symptoms.

Lehmkuhl et al., (2008) document similar modifications and focused on a behavioural mechanism of change i.e. ERP, rather than cognitive techniques. They also adapted abstract concepts to tailor to the individual's developmental stage. After ten sessions of CBT, based on March and Mulle's (1998) protocol also, the young person similarly

achieved remission status. The only RCT to date focusing on treating specifically OCD in ASD recruited a wide ranging age of participants, aged 14-65 years (Russell, et al., 2013). The findings suggest that both anxiety management and CBT show statistically equivalent and meaningful response rates (20% vs. 45%, respectively) and mean reductions in Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman et al., 1989) scores at the end of treatment, although the effect size for CBT was larger than for AM (1.15 vs. .6, respectively). Similarly, remission rates were statistically equivalent between CBT and AM (20% vs. 15%). The treatment gains during the acute phase of CBT were maintained over twelve months in the CBT condition. This provides very promising evidence for the use of CBT and potentially other treatment modalities in treating OCD in adolescents and adults with an ASD.

The limitation of generalizability is clear in the case studies and to a paediatric population in the predominantly adult RCT, but it is also worth noting that while both case studies fully remitted by the end of treatment, their initial OCD severity scores as measured by the CY-BOCS were only within the mild (Lehmkuhl, et al., 2008) and moderate ranges (Reaven & Hepburn, 2003). Similarly, in Russell et al.,'s (2013) trial the mean initial CY-BOCS for both treatment groups was in the moderate range. It is well attested that severity of OCD symptomatology affects treatment outcome (Ginsburg, Kingery, Drake, & Grados, 2008), which questions how effective CBT treatments are for individuals who have a co-morbid ASD and OCD presentation, where their OCD is more severe?

1.7 Service Evaluation Aims and Hypotheses

OCD is relatively common in young people with ASD and while evidence is accumulating for CBT's effectiveness using modified protocols for other anxiety disorders, the evidence for specifically treating OCD in young people with an ASD with CBT is very limited. In light of the above, the service evaluation had three main aims:

Aim 1) To identify the rates of young people being referred with a dual diagnosis of ASD and OCD (ASD+OCD) to the National Specialist OCD service, within which this study was conducted.

Aim 2) To assess the initial demographic and clinical characteristics of this cohort when compared to referrals who received an OCD diagnosis, but who did not have an ASD (NoASD+OCD).

Aim 3) To assess how this cohort responded to a standard protocol of CBT for OCD when compared to individually matched and typically developing NoASD+OCD controls.

Three hypotheses were generated:

Hypothesis 1) Rates of referrals of the ASD+OCD subgroup have increased over the last five years,

Hypothesis 2) The ASD+OCD subgroup's initial OCD symptomology would be more severe as compared to the NoASD+OCD group.

Hypothesis 3) The ASD+OCD subgroup would have responded less well to a course of CBT for OCD compared to an age, gender and OCD severity matched NoASD+OCD group.

1.8 Service Context

1.8.1 The National Specialist OCD and Related Disorders Service

The National Specialist OCD and Related Disorders service at the Maudsley Hospital is a Tier 4 outpatient service that receives referrals nationwide, but which mainly receives referrals from local Tier 3 Children and Adolescent Mental Health Teams (CAMHS). While the majority of referrals are for young people affected by OCD, disorders that are related or likely to be co-morbid e.g. Body Dysmorphic Disorder, Tourette's Syndrome and Trichotillomania, are also seen within the service. The team is formed of experienced Psychologists and Psychiatrists and offers both pharmacological and psychological treatments. Psychological treatments are delivered either on a weekly basis face to face/ over the phone or over the course of a 5 day intensive program where therapists work all day with families either at the clinic or at the patient's family home.

Assessment and treatment funding is usually commissioned through local commissioning groups although a number of referrals are funded via the National

Specialist Commissioning Team (NSCT). For patients to meet NSCT status their OCD must be within the severe range as measured by well validated structured assessments. Secondly, the individual must have had an ineffectual trial of a serotonin reuptake inhibitor (SSRI) at a recommended dose for at least three months and finally, the patient must have had a course of CBT that includes ERP that has also been ineffectual in reducing OCD symptomatology.

1.8.2 Assessment Process

All patients who are seen in the service undergo a thorough multidisciplinary assessment process that is overseen by a consultant Psychiatrist or Psychologist. The whole process lasts approximately three hours and during which, both parent interviews, focusing on a developmental history of the young person and important clinical information to inform ICD-10 diagnosis are conducted alongside a child interview focusing on the young person's OCD symptomatology and severity. Diagnoses were assigned or confirmed at the point of assessment in the OCD Clinic by the multi-disciplinary team. With respect to ASD diagnoses, the majority of cases were referred with the diagnosis made in Tier 3 CAMHS. The remainders were referred without a diagnosis, but it was assigned by the OCD clinical team following a structured diagnostic assessment such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989)

1.8.3 Cognitive Behavioural Therapy Protocol

The CBT protocol follows three distinct stages. In the first stage, sessions are focused on psychoeducation about OCD and anxiety, goal setting and hierarchies of the young person's OCD symptoms are developed. In the second stage, which forms the majority of treatment, sessions are focused on ERP and collaboratively working up the young person's 'fighting back' hierarchy. The final stage is orientated towards relapse prevention. Parents are involved where possible and where clinically appropriate e.g. if family members are accommodating OCD rituals, but as a minimum, parents and family members are involved during the psychoeducation and relapse prevention stages.

2 Method

2.1 Participants

Consecutive referrals (N=387) to the OCD clinic that were assessed between January 2007 and December 2011 were reviewed. For the purpose of examining referral rates, the dataset was categorized into 4 sub-groups: 1) Individuals who met diagnostic criteria for both OCD and an ASD at assessment (ASD+OCD); 2) Individuals who met diagnostic criteria for OCD, but did not meet criteria for an ASD at assessment (NoASD+OCD); 3) Individuals who met diagnostic criteria for an ASD, but did not meet diagnostic criteria for OCD at assessment (ASD+NoOCD); and 4) Individuals who did not meet diagnostic criteria for either an ASD or OCD (NoASD+NoOCD).

A sub-group of the ASD+OCD group received CBT in the OCD clinic and had outcome measures available (n=22). These were compared against a group of young people within the OCD group (n=22)

2.2 Measures

2.2.1 The Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS)

The CY-BOCS (Scahill, et al., 1997) is the gold standard measure for assessing OCD symptom severity in children and is based on the adult Y-BOCS (Goodman, et al., 1989). For this reason it was considered the primary outcome of interest. The CY-BOCS is an interviewer rated, semi-structured clinical interview. It comprises of 5 items (time spent, interference, distress, resistance and control) for both obsessions and compulsions, from which a total OCD severity score ranging from 0-40 can be calculated. The CY-BOCS shows good inter-rater reliability, construct validity and has been validated within a paediatric ASD population (Wu et al., 2013)

2.2.2 The Child Obsessive Compulsive Inventory-Revised (ChOCI-R)

The ChOCI-R (Shafran et al., 2003) is an OCD symptom severity questionnaire that follows the same format as the CY-BOCS and adult version of the Obsessive Compulsive

Inventory (Foa, Kozak, Salkovskis, Coles, & Amir, 1998). There are both a young person and a reliable informant version, typically a parent, and the measure shows good internal consistency and convergent validity (Uher, et al., 2008)

2.2.3 The Children's Global Assessment Scale (CGAS)

The CGAS (Shaffer et al., 1983) was used to measure overall clinical severity and is a well validated measure (Shaffer, et al., 1983). Scores range from 0 to 100 with scores above 70 representing normal functioning. The CGAS shows good inter-rater reliability and shows good concurrent and discriminant validity (Shaffer, et al., 1983)

2.3 Information Obtained

Demographic information that was systematically extracted from case notes included:

- Age in years, gender, age of OCD onset and NSCT status.

The following clinical Information was obtained for individuals within the treatment outcome analysis:

- Session number, medication status and whether medication and CBT started concurrently.

2.4 Procedure

Data were extracted from an existing electronic database. In cases where key variables of interest were missing, attempts were made to retrieve data by examining paper and electronic files. This data was then examined and entered onto the clinic's routine assessment and outcome SPSS database.

In order to examine treatment outcomes of young people with ASD+OCD, a matched subgroup from the NoASD +OCD group were identified. CY-BOCS score at assessment was used as the primary matching variable as baseline OCD severity has been shown to affect CBT treatment outcomes (Ginsburg, et al., 2008). In addition, the groups were matched for age and gender. Approval to conduct the audit was granted by the South London and Maudsley Child and Adolescent Mental Health Service Audit Committee.

3 Results

3.1 Referral rates

Figure 1: Frequency of referrals by group between 2007 and 2011.

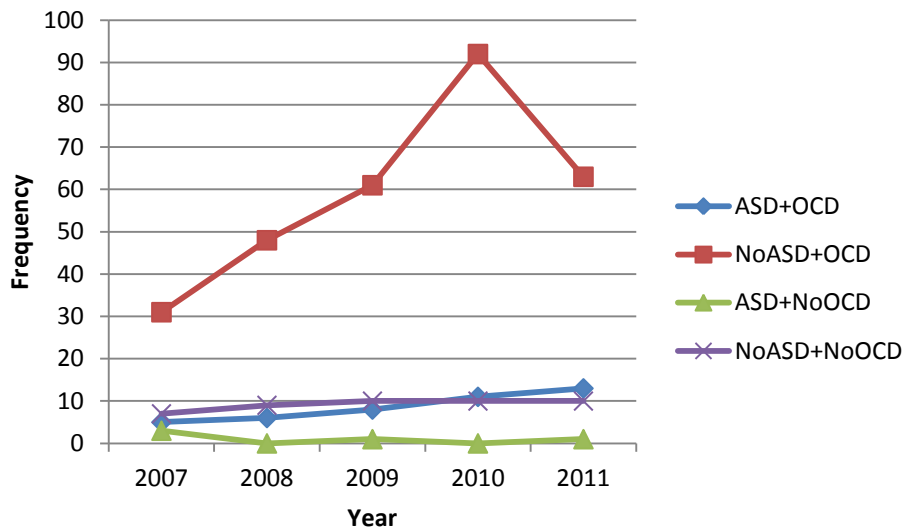
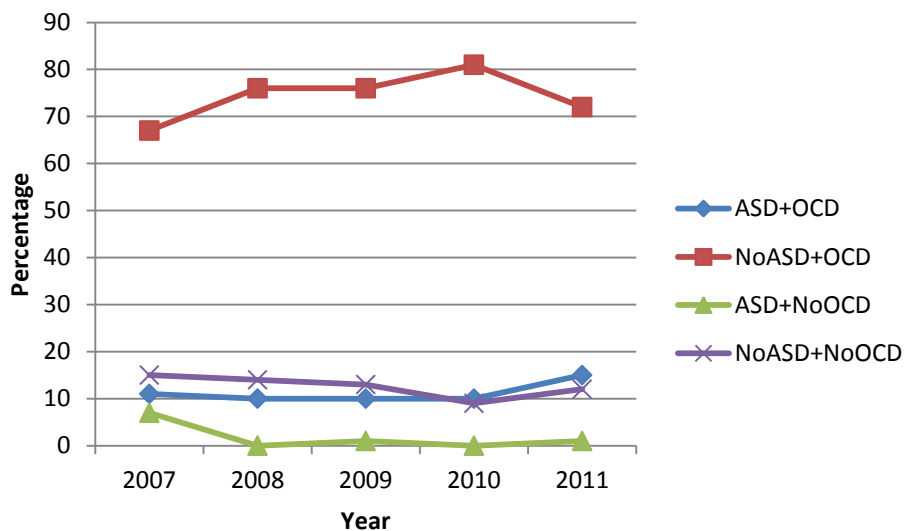


Figure 2: Relative percentage of each group when compared to the overall number of referrals per year between 2007 and 2011.

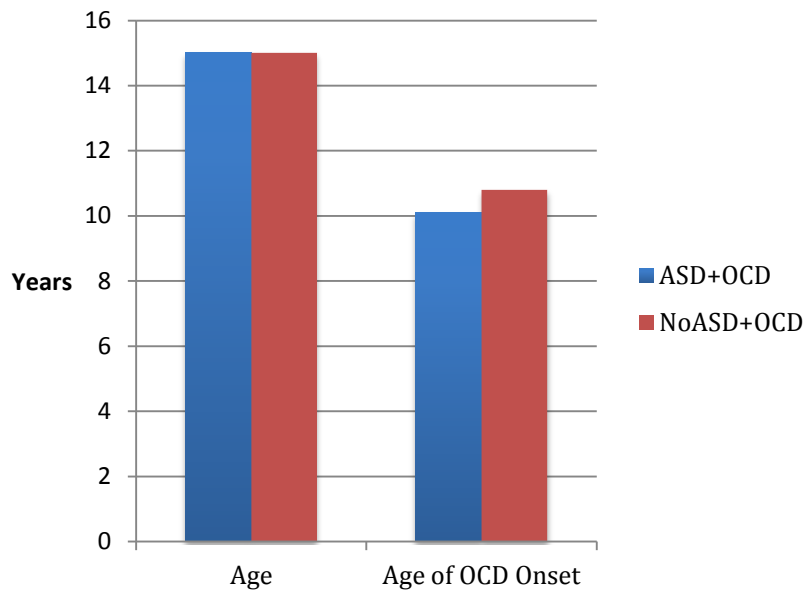
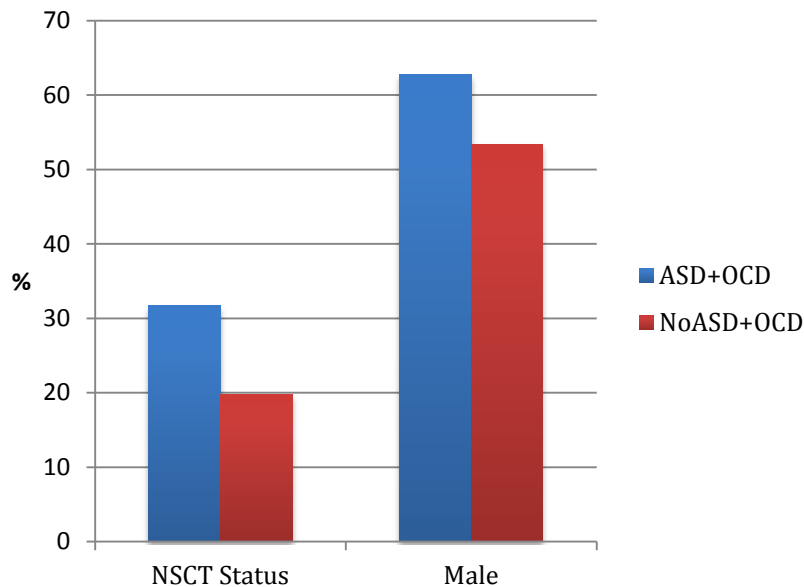


Frequencies of referral rates of each sub-group across the five year period are summarized in Figure 1 (above) and Figure 2 shows the data as percentages of the overall

referrals for that year. The ASD+OCD group's referral rates increased across the five year time period from 5 to 13 although as a percentage of the clinic's referrals the referral rate was stable from 2007-2010 accounting for 10% of referrals, with a slight increase in 2011 to 15% of referrals. Rates of referral for the NoASD +OCD group increased steadily from 31 in 2007 to 92 in 2010, but then declined to 63 in 2011. Similarly, as a proportion of the overall clinic's referrals the referral rates for the NoASD +OCD group increased from 67-81% between 2007 and 2010 and then dropped to 72% in 2011. As expected given the service context, the number of ASD+NoOCD referrals was consistently low across the time period being studied, with the number of referrals decreasing from 3 (7%) in 2007 to either 0 or 1 between 2008 and 2011. Finally, rates of referral for the NoASD+NoOCD group remained consistent throughout the time period, with referrals ranging from 7-10 per year (9-15%), across the five years.

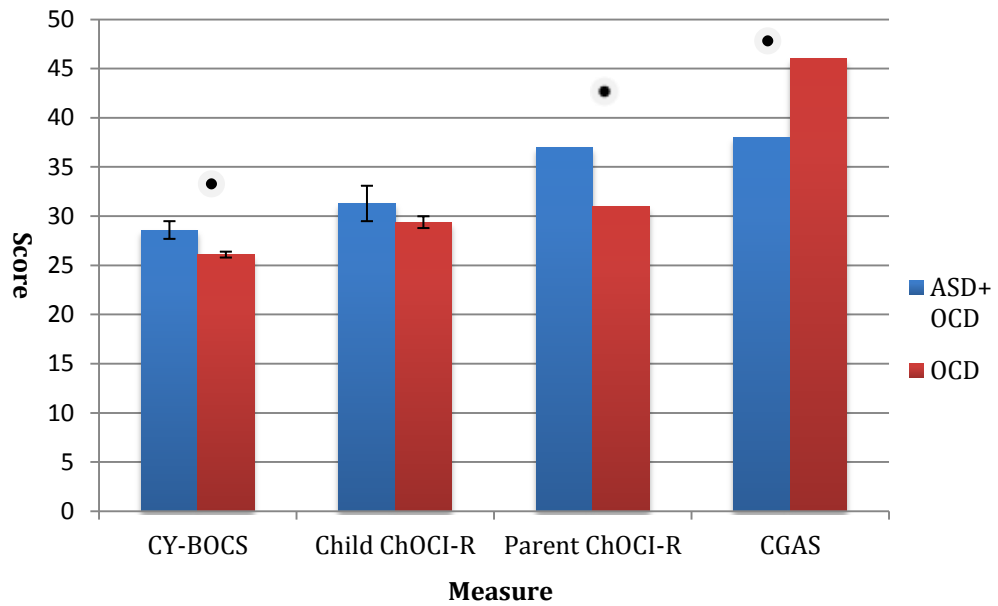
3.2 Demographics and Clinical Characteristics at Assessment

The ASD+OCD and the NoASD +OCD group were compared on key demographic and clinical variables. Only individuals who had completed a CY-BOCS interview were included in this section of the analysis (ASD+OCD group: N= 37; NoASD+OCD group: N= 283). Parametric properties were assessed for all variables and in cases where parametric assumptions were violated (e.g. data's variance was not homogenous) non-parametric t-test equivalents were used. In all figures therefore, means are depicted with error bars in cases where assumptions were met, and in all other cases medians are presented. Comparisons were made between CY-BOCS scores for each of the other routine measures between individuals who had completed the measure and those who had not, which revealed no significant differences (all p-values > 0.05).

Figure 3: Age of both groups and age of OCD onset in years**Figure 4: Percentage of each group meeting NSCT funding criteria and the percentage of males in each group.**

There was no difference between the ASD+OCD and NoASD+OCD groups with respect to age at assessment, $U = 5260$, $z = -1.41$, $p = .16$, or age at onset of OCD, $t(306) = -1.10$, $p = .27$. The ASD+OCD group had slightly more males and individuals referred under NSCT status than the NoASD+OCD this did not reach statistical significance, $\chi^2(1, 326) = 1.34$, $p = .25$ and $\chi^2(1, 309) = 3.014$, $p = .08$, respectively.

Figure 5: Total CY-BOCS, Child and Parent ChOCI-R, and CGAS scores at assessment.



• $p < .05$.

Analysis between groups yielded a significant difference between CY-BOCS scores, $t(318) = 2.50$, $p = .013$ with the ASD+OCD initial score ($M = 28.6$, $SD = 5.3$) being larger than the NoASD+OCD group ($M = 26.1$, $SD = 5.7$). Scores on the Parent ChOCI were also significantly larger, $U = 2432$, $z = -2.78$, $p = .005$, within the ASD+OCD group ($Mdn = 37$, $IQR = 12$) compared to the OCD group ($Mdn = 31$, $IQR = 15$). However, there was no significant group difference with respect to scores on the Child ChOCI $t(253) = 1.117$, $p = .27$. Finally, CGAS scores were significantly lower, $U = 1665$, $z = -4.40$, $p = .001$, in the ASD+OCD group ($Mdn = 38$, $IQR = 10$) compared to the NoASD+OCD group ($Mdn = 46$, $IQR = 14$).

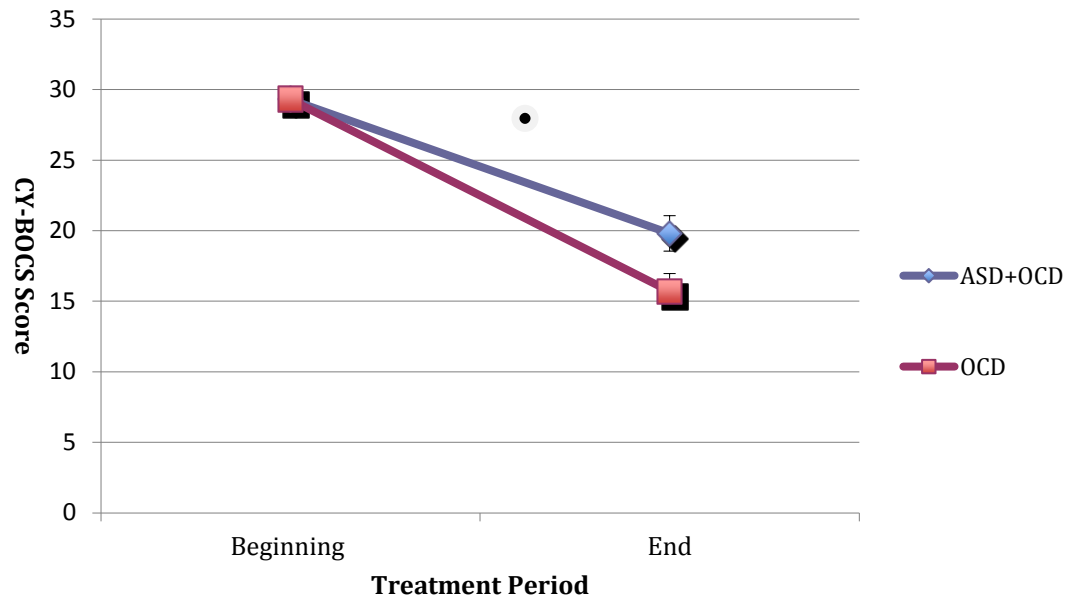
3.3 Treatment

All individuals within the ASD+OCD group that had also completed a course of CBT within the service were identified (N=22). The ASD diagnoses of these individuals were as follows: Asperger's Syndrome (N=15), High Functioning Autism (N=2) and PDD-NOS(N=5). 68% of the group had a supporting ADI-R or ADOS. All of the individuals in this group had an IQ in the normal range.

Patients in the ASD+OCD group were then individually matched to individuals from the NoASD+OCD group using CY-BOCS score at baseline as the primary matching variable, followed by age and gender (see table 2). A repeated measures mixed ANOVA was run with a within subjects factor of treatment (Beginning vs. End) and a between subjects factor of group (ASD+OCD vs. NoASD+OCD). In addition, CBT treatment response rates (CY-BOCS score reduction of at least 25%) and symptom remission rates (end of treatment CY-BOCS score of 14 or less) were calculated using empirically derived cut-offs (Storch, Lewin, De Nadai, & Murphy, 2010) as a measure of treatment efficacy. Rates of stable prescribed selective serotonin re-uptake inhibitors (SSRI's) were calculated between the groups alongside whether individuals in the group had been started on an SSRI alongside CBT treatment. Finally, the number of CBT sessions was analyzed between the groups.

Table 13: Matched Group's Initial Presentation

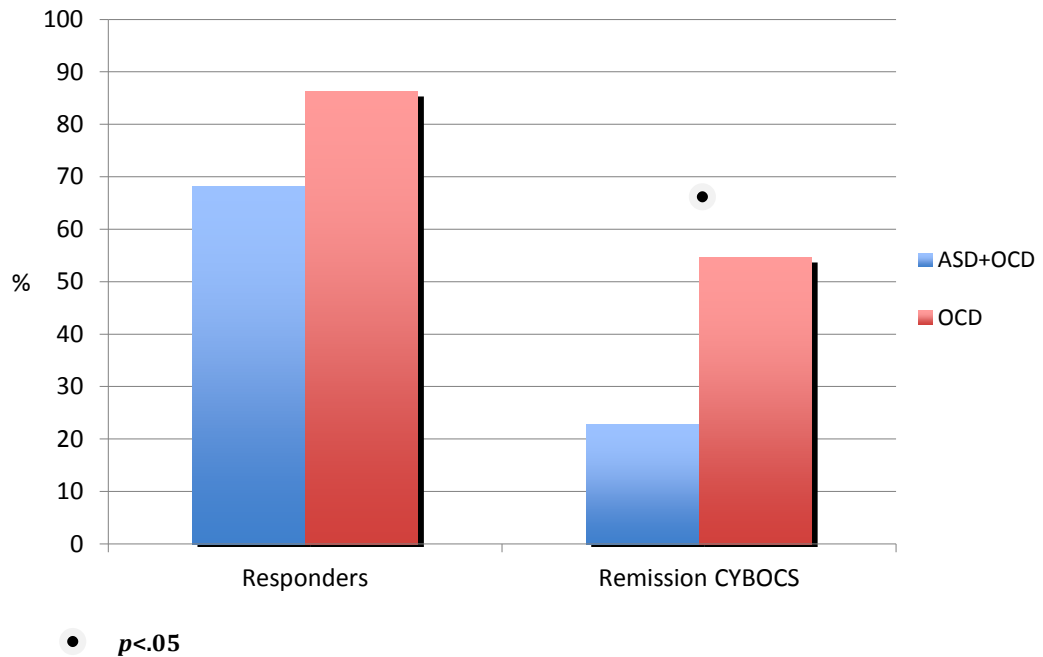
| | Group | |
|-----------------------------|--------------|--------------|
| | ASD+OCD | NoASD+OCD |
| N | 22 | 22 |
| CY-BOCS Score <i>M</i> (SD) | 29.36 (4.98) | 29.36 (5.01) |
| Age <i>Mdn</i> (IQR) | 15 (3.25) | 15(4.25) |
| Gender (males: females) | 14:8 | 12:10 |

Figure 6: Mean total CY-BOCS score at the Beginning and End of treatment.

● $p < .05$.

Analysis revealed that there was a significant main effect of treatment for both groups $F(1,42) = 177.699$, $p = .001$ and a significant interaction between group and treatment $F(1, 42) = 5.69$, $p = .022$. The ASD group's CY-BOCS mean difference ($M = -9.59$, $SD = 4.91$, $D = -1.95$) was significantly lower when compared to the OCD group's mean difference ($M = -13.77$, $SD = 6.59$, $D = -2.09$) at the end of treatment. The interaction's effect size was $D = .72$.

Figure7 : Percentage of treatment responders (Responders) and remission rates as measured by CY-BOCS score of 14 or less (Remission CY-BOCs).



Analysis revealed no significant effect of group when comparing the number of treatment responders ($X^2 (1, 44) = 2.10, p = .15$). However, a statistical effect was revealed between remission rates as measured by a CY-BOCS of 14 or less at the end of treatment, $X^2 (1, 44) = 4.7, p = .03$, where the ASD group showed lower rates of remission (23%) compared to the OCD group (55%).

Furthermore, analysis showed that the likelihood of a participant being on a stable dose of an SSRI did not differ significantly between the two groups although there was a trend towards the ASD group being more likely to be on an SSRI ($X^2 (1, 44) = 3.30, p = .07$). Analysis revealed that group membership did not affect the likelihood of starting pharmacotherapy and CBT concurrently ($X^2 (1, 44) = .61, p = .43$). While session numbers were statistically equivalent between the two groups there was a trend toward the ASD+OCD group having more sessions than the NoASD+OCD group ($U = 287, z = -1.87, p = .06$).

4 Discussion

4.1 Referral Rates

The first aim of this audit was to examine referral rates of young people with a dual diagnosis of an ASD and OCD, which have increased from 2007-2009 although this increase is only minimal. However, when compared to a previous paper examining rates of referrals within the same clinic (Mack, et al., 2010) the referrals have increased substantially from 3.8% of referrals between 1996-2006 to currently the subgroup representing 15% of referrals in 2011. This increase in referral rate may reflect a number of factors, including the increasing profile of the service, and the introduction of the NSCT funded branch of the service, for severe, treatment-resistant OCD. Importantly, the rates of ASD referrals without OCD being referred was nearly zero throughout the time period suggesting that referrers are able to distinguish when OCD is not present. Importantly this does not equate to accurately discriminating OCD from ASD symptomatology when OCD is present.

Overall, this points to an ever increasing need for accurate assessment within the clinic as this subgroup now represents a substantial proportion of referrals. Screening of ASD symptomatology where appropriate could aid this process. The Social Communication Questionnaire (SCQ) is a well validated screening tool of ASD symptomatology in young people that is not affected by parental IQ or education (Chandler et al., 2007). Using such screening tools would provide additional information to clinicians working with this group without the burden of more intensive diagnostic tools such as the ADOS (Lord, et al., 2000), which could be used for more select cases.

Identifying and subsequent differentiation of symptoms early has benefits in treatment outcome as although a review of predictors of treatment response to both psychological and pharmacological treatments of paediatric OCD (Ginsburg, et al., 2008), showed that duration of illness did not affect response rates, long-term follow-up data suggests that duration of illness at assessment is a predictor of having a diagnosis of OCD 5 years after treatment (Micali et al., 2010). While this data set did not include those with an ASD it is possible that such individuals could be similarly more likely to relapse if they have had OCD symptoms for a longer period prior to assessment and subsequent treatment.

4.2 Demographics and Initial Presentation

The second aim was further understand the demographic and clinical variables of young people with a dual ASD and OCD diagnosis. The demographic findings suggest that the age at onset of OCD is similar among those with and without ASD, they are referred to the clinic at a similar age, and are no more likely to be referred as part of the National Specialist Commissioning Team. However, the ASD+OCD group had significantly more severe OCD symptoms as rated by clinicians and parents, but not by the young people being referred. Additionally their overall functioning was rated as being lower as rated by clinicians.

One interpretation of these findings would suggest that the construct validity of self report measures assessing OCD in young people with an ASD is undermined, which adds evidence to one side of Wood and Gadow's (2010) discussion on this area. This poses the additional question of whether self-report questionnaires should be used at all in young people with ASD. While the question is rhetorical, the utility of self-report and its function in the clinical assessment of OCD in ASDs is an important issue to clarify. Additional clinical information can clearly be obtained from self report measures, such as the content of obsessions, which could possibly be less available to informants (Uher, et al., 2008) or understanding which part of the young person's OCD they articulate as most distressing. The latter is particularly helpful in engaging young people in the therapeutic process and setting personally meaningful goals.

The discrepancy between self and clinician/parent report could also be used to help guide the focus of psychoeducation, such as exploring symptoms the young person may be less aware of using developmentally appropriate language and concrete visual aids (Reaven & Hepburn, 2003). Self-report instruments might be more adept at aiding the aforementioned clinical decisions therefore rather than guiding decisions regarding overall OCD severity in this population. The study is clearly limited with regard to understanding the potential confounding variables that may be affecting the observed discrepancy such as intra group differences in alexithymia symptoms (Tani, et al., 2004) or IQ. This is an area that warrants further research. Conversely, the 2 point difference observed between the groups means could be argued as representing a statistical, but not clinically meaningful finding. The weight of clinical experience within the National

Specialist OCD clinic when assessing and treating co-morbid OCD within an ASD population (Russell, et al., 2013) and the construct validity and treatment sensitivity of the CY-BOCS in this population (Wu, et al., 2013) should support the notion that these findings are not related to OCD symptoms being missed or misattributed to being part of the ASD phenotype. In all, further research is needed to inform what patient, informant and clinician variables are moderating this process. Clinically, this finding necessitates the additional input of multiple informants when conducting psychiatric assessments with youth with an ASD and OCD. It is only with such corroborative evidence that a meaningful clinical description of the young person's OCD be reached.

4.3 Treatment Outcomes

The third and final aim of this audit was to evaluate treatment outcomes for young people with a dual diagnosis of ASD and OCD compared to a case matched controlled sample. Following a course of evidence-based treatment by a team of experienced therapists who work with both severe OCD in youth and OCD in the context of an ASD (Russell, et al., 2013) the ASD+OCD group's mean difference was significantly lower than the matched OCD group at the end of treatment. The ASD+OCD group fell within the 'Moderate' range of severity and the OCD group fell within the 'Mild' range of severity (Scahill, et al., 1997). The effect size of the interaction was medium ($d=.72$) and the effect of treatment for both the ASD+OCD and OCD group was large ($d=-1.95$ and $d=-2.09$, respectively). Importantly, the differences in CBT response found between the two groups cannot be explained in terms of group difference in medication or number of CBT sessions.

The pattern of response rates supports this picture further. No differences were revealed between response rates, which supports the growing body of literature that CBT is effective at treating anxiety disorders (Lang, et al., 2010; Reaven, et al., 2012; Storch, et al., 2013) and OCD specifically (Lehmkuhl, et al., 2008; Reaven & Hepburn, 2003; Russell, et al., 2013) in youth with an ASD. It is worth noting that the response rates observed are similar to some previous RCTs (Lang, et al., 2010; Storch, et al., 2013), but higher than others (Russell, et al., 2013). The former trials, with comparable response rates, were solely working with young people there may be an effect of age occurring in that overall young people with ASD respond better to treatment than their adult counterparts. It is possible that the greater levels of involvement from family members and the young

person's wider system may be contributing to this effect (Reaven, et al., 2012) and that for adults with ASD they may also benefit from such input if it is available, to help with difficulties characteristic of having an ASD e.g. generalizing gains made during therapeutic time.

Remission rates were significantly lower for the ASD+OCD group compared to the OCD group (23% vs. 55% respectively). The remission rates are similar to previous findings (Russell, et al., 2013) but lower than other RCTs which included young people with ASD and anxiety disorders (Storch, et al., 2013). It is worth noting that the use of differing definitions of remission, outcome measures and varying levels of baseline severity between studies will contribute to these inconsistencies especially as response is a relative definition whereas remission is an absolute term. Furthermore, previous studies would have employed an adapted CBT protocol tailored specifically to the needs of young people with ASD, whereas in the current study a standard CBT protocol was adopted.

The pattern of results suggests that there are two key points. Firstly, that baseline OCD severity is highly influential in how well the individual will respond to CBT treatment akin to youth without an ASD (Ginsburg, et al., 2008) and this is naturally going to affect clinical remission status. The second is that when these two presentations occur in unison, severe OCD and ASD, the individual's OCD is particularly difficult to ameliorate altogether. The overall severity of the group was therefore an obstacle in itself to treatment, but not contraindicative that psychological therapy could be beneficial for this population.

What can be taken from this evidence and the research precluding it, is that individuals with an ASD and co-morbid OCD benefit from CBT, which puts to rest the notion that this population does not have the prerequisite skills necessary to engage with this form of therapy (for further discussion on this topic see Lickel, MacLean, Blakeley-Smith, & Hepburn, 2012), but improving remission rates is a challenge for the research and clinical communities. This is an important consideration as improved remission rates not only benefit young people in the short term with regard to relief from the distress associated with the condition, but also previous trials have showed that young people who reach such status maintain their gains at follow-up (Russell, et al., 2013; Storch, et al., 2013). It is clear that a standard CBT protocol is not sufficient for this population and that modification's are key to improving clinical outcomes for this client group.

4.4 Treatment Modifications

While it is beyond the scope of this service evaluation to discuss all the treatment modifications that would be helpful in this client group and further literature is available on this area a number of themes that are present in the literature are evident. Firstly, this client group requires a greater number of sessions to benefit from CBT (Russell, et al., 2013). The family and broader system may need to be involved more heavily and not just within the psychoeducation and relapse prevention phases (Reaven, et al., 2012) to help with clinical features of ASD e.g. generation difficulties and to help parents delineate ASD symptomatology from OCD symptoms. Clinicians may have to alter their interactional style to account for interpersonal, imagination difficulties and executive function difficulties such as planning associated with the diagnosis. Therapists may want to spend more time and focus on psychoeducation especially with regard to delineating thoughts, feelings and behaviours and how they influence on another and exploring/ identifying emotions as alexithymia has been shown to be elevated in this population (Tani, et al., 2004).

When completing treatment a number of adaptations surrounding making E/RP 'autism friendly' is helpful e.g. using an individual's special interest to complete an anxiety scale and using more tangible signs of anxiety e.g. physical symptoms, as a gauge on how anxious the young person is feeling rather than relying on a scale of one to ten where the question used is "How worried do you feel right now?". Similarly, Reaven & Hepburn (2003) highlight the importance of adapting the language used with this client group. This could include avoiding complex metaphors that involve multiple perspectives, and using concrete visual aids during E/RP to promote the young person's abilities to utilize their own resources when facing anxiety provoking situations. The focus of the treatment may therefore need to be more behavioural in nature (Lehmkuhl, et al., 2008). Recent research has shown that treating anxiety can have a positive effect on the young person's ASD symptomatology as well as their anxiety symptoms (Wood, et al., 2009) suggesting that it may be beneficial to broaden the focus of treatment and include outcome measures that can evidence this if it occurs. In all, the intervention needs to be 'neurodevelopmentally' friendly.

4.5 Limitations

The current study has a number of limitations. First of all, therapists would have acted as raters for treatment outcome potentially biasing the main treatment effect sizes. However, while this may have altered the neutrality of the findings this effect is stable across both groups so does not undermine the interaction. Secondly, the discrepancy in session numbers (although not significant) means that gains made after 14 sessions within 17 weeks as per the clinic protocol cannot be assessed. Thus, the high response rate in the ASD group could be due to increased therapeutic contact. While this is a limitation in the findings, it provides further evidence that youth with ASD and OCD require a greater length of therapeutic involvement from clinicians to benefit from CBT delivered by experts in the field. Previous research protocols have involved CBT interventions between 16 and 20 sessions in length (Russell, et al., 2013; Storch, et al., 2013) as opposed to 14 seen in RCTs for typically developing youth (POTS, 2004) as is the norm for the clinic. It is important in the absence of clinical guidelines for commissioners that funding bodies are aware of this when negotiating treatment lengths that are going to be effective for this client group.

Additional co-morbid diagnoses were not systematically screened for. In future studies it would be important to consider this closer, as depression, attention deficit hyperactivity disorder (ADHD) and opposition defiant disorder (ODD) all impact on response and/or remission rates. Similarly, having at least one or more co-morbid psychiatric condition also impacted on end of treatment outcome (Storch et al., 2008). Thus, systematic screening of these diagnoses would be particularly helpful in improving outcomes and further research is needed to assess whether this is true in the case of young people with an ASD as well.

Both groups were statistically equivocal in terms of whether individuals were taking an SSRI or not and whether individuals started both pharmacological and CBT treatments concurrently. However, as some individuals were not on stable medications during the treatment phase, teasing apart the relative contributions of CBT and the SSRI on treatment outcome is difficult.

Finally in the absence of follow up data it is unclear whether the gains made during the acute phase of treatment would be maintained over a follow-up period. The adult data suggests that this population is able to maintain gains it's gains, but that OCD symptoms

plateau over a follow-up period (Russell, et al., 2013) and participants do not continue to improve and reach remission status. The short term data in youth similarly show no additional gains over a short term follow up period of three months (Storch, et al., 2013). Further research is needed to clarify a) whether there is a differential in terms of follow-up trajectory and b) how treatment can be adapted if a differential exists.

5 Service Recommendations

In this section a number of service recommendations will be suggested that may help the service manage their ASD+OCD referrals on the basis of the study's findings:

- **Systematically implementing a well validated screening tool e.g. SCQ (Chandler, et al., 2007) for referrals who are likely to have an ASD, prior to assessment and at the end of treatment.**
- **Ensure information from multiple informants' e.g. teachers, parents and clinicians when it is available and spending additional time to obtain evidence from such sources if the information is not readily available.**
- **Prioritize clinician rated measures e.g. the CY-BOCS, which have been validated in an ASD population to inform OCD severity in the assessment process.**
- **Use the assessment process as part of the intervention to disentangle ASD from OCD symptoms e.g. assessing the presence of obsessions and identifying egodystonic beliefs regarding said symptoms.**
- **Systematically assess for the presence of co-morbid psychiatric disorders especially ADHD, ODD and depression, which can impact on treatment outcome.**
- **Discuss with funders about negotiating additional treatment time of up to 20 sessions for the treatment phase.**
- **Draw upon treatment manuals in the ASD+OCD literature to develop a paediatric ASD+OCD specific intervention so that the team feels confident in treating such cases and a consistent protocol driven approach is adopted akin to typically developing referrals. Peer supervision and/or training within the team may be helpful.**
- **Once a modified CBT protocol for ASD+OCD has been developed, this should be evaluated in controlled studies.**

6 Dissemination of Results

The results of this service evaluation were presented to the OCD team on their research 'away day.' In addition, as part of a symposium chaired by Dr Georgina Krebs, on 'obstacles in treating paediatric OCD', this study was presented at the BABCP conference in London, 2013. The findings been submitted to the European Journal of Child and Adolescent Psychiatry and is currently under peer review. The findings were used to support the application of a grant for a randomized control trial investigating the treatment efficacy of modified CBT in young people with ASD and OCD.

7 Reflections

This service evaluation has provided a space for learning about the intimate connection between academic understanding and the development of service provision within SLAM. From preparing for ethical approval, collating and analyzing the data and finally disseminating the findings to the team and the wider academic community I have learned that to contribute to the development of a service provision it is crucial to look at the individuals who the current forms of treatment are less efficacious and from there one can better a service and the treatments it provides.

This work has given me insights into obtaining routine outcome measures and how valuable the data is for bettering services more generally. Without the diligence of the team to manage their clinical data efficiently, research like this would not be possible and it is only with such data that the evidence-base can be evaluated in clinical settings outside the academic arena. When considering the political and economic climate of the last five years maintaining and evidencing a services work through robust and well validated outcome measures is all the more important.

Furthermore, this work has highlighted the need for further research in this area as the number of well designed studies evaluating treatment efficacy in young people with ASD and anxiety disorders is sparse and it is gratifying to think that this piece of work may contribute to the clinical and academic communities understanding of working with young people with these conditions.

Presenting the work was a challenge and as it was the first time I have presented at a major conference so it was naturally anxiety provoking. However, the simultaneous satisfaction of sharing a piece of work with an interested and professional audience was extremely rewarding. Their questions, comments and reflections on the work guided my thinking about not only this work, but the importance of academic and clinical practice evolving simultaneously.

With hindsight it would have been helpful to know more about how the young people's ASD symptoms interacted with treatment outcome. Furthermore, analyzing the effects of treatment on the young people's obsessions and compulsions separately would have been

interesting to have understood. However, this work has taught me that research is slow growing and it is important to understand complex phenomena such as this piece by piece.

8 References

Albano, A. M., Chorpita, B. F., & Barlow, D. H. (2003). Childhood anxiety disorders. In E. J. Mash & R. A. Barkley (Eds.), *Child Psychopathology* (Vol. 2, pp. 279-329). New York: Guilford Press.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, 368(9531), 210-215.

Baron-Cohen, S. (1989). Do autistic children have obsessions and compulsions? *The British Journal of Clinical Psychology*, 28(3), 193-200.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind” ? *Cognition*, 21(1), 37-46.

Chalfant, A. M., Rapee, R., & Carroll, L. (2007). Treating anxiety disorders in children with high functioning autism spectrum disorders: a controlled trial. *Journal of Autism and Developmental Disorders*, 37(10), 1842-1857.

Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T., Meldrum, D., et al. (2007). Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(10), 1324-1332.

de Bruin, E. I., Ferdinand, R. F., Meester, S., de Nijs, P. F. A., & Verheij, F. (2007). High rates of psychiatric co-morbidity in PDD-NOS. *Journal of Autism and Developmental Disorders*, 37(5), 877-886.

Department of Health Payment by Results Team. (2012). A simple guide to Payment by Results.18135. Leeds: Author.

Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a new obsessive-compulsive disorder scale: The Obsessive-Compulsive Inventory. *Psychological Assessment*, 10(3), 206-214.

Gadow, K. D., Devincent, C. J., Pomeroy, J., & Azizian, A. (2005). Comparison of DSM-IV symptoms in elementary school-age children with PDD versus clinic and community samples. *Autism*, 9(4), 392-415.

Geller, D., Biederman, J., Jones, J., Shapiro, S., Schwartz, S., & Park, K. S. (1998). Obsessive-compulsive disorder in children and adolescents: a review. *Harvard Review of Psychiatry*, 5(5), 260-273.

Ghaziuddin, M. (2002). Asperger Syndrome Associated Psychiatric and Medical Conditions. *Focus on Autism and Other Developmental Disabilities*, 17(3), 138-144.

Ginsburg, G. S., Kingery, J. N., Drake, K. L., & Grados, M. A. (2008). Predictors of Treatment Response in Pediatric Obsessive-Compulsive Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(8), 868-878.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., et al. (1989). The Yale-Brown obsessive compulsive scale: I. Development, use, and reliability. *Archives of General Psychiatry*, 46(11), 1006-1011.

Heyman, I., Mataix-Cols, D., & Fineberg, N. A. (2006). Obsessive-compulsive disorder. *BMJ: British Medical Journal*, 333(7565), 424.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, 8(1), 26-32.

Lang, R., Regeister, A., Lauderdale, S., Ashbaugh, K., & Haring, A. (2010). Treatment of anxiety in autism spectrum disorders using cognitive behaviour therapy: A systematic review. *Developmental Neurorehabilitation*, 13(1), 53-63.

Lehmkuhl, H. D., Storch, E. A., Bodfish, J. W., & Geffken, G. R. (2008). Brief report: exposure and response prevention for obsessive compulsive disorder in a 12-year-old with autism. *Journal of Autism and Developmental Disorders*, 38(5), 977-981.

Lewin, A. B., Wood, J. J., Gunderson, S., Murphy, T. K., & Storch, E. A. (2011). Phenomenology of comorbid Autism Spectrum and obsessive-compulsive disorders among children. *Journal of Developmental and Physical Disabilities*, 23(6), 543-553.

Lickel, A., MacLean, W. E., Blakeley-Smith, A., & Hepburn, S. (2012). Assessment of the prerequisite skills for cognitive behavioral therapy in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42(6), 992-1000.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205-223.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, 19(2), 185-212.

Mack, H., Fullana, M. A., Russell, A. J., Mataix-Cols, D., Nakatani, E., & Heyman, I. (2010). Obsessions and compulsions in children with Asperger's syndrome or high-functioning autism: a case-control study. *The Australian and New Zealand Journal of Psychiatry*, 44(12), 1082-1088.

March, J. S., & Mulle, K. (1998). OCD in children and adolescents: A cognitive-behavioral treatment manual. New York: Guilford Press.

Micali, N., Heyman, I., Perez, M., Hilton, K., Nakatani, E., Turner, C., et al. (2010). Long-term outcomes of obsessive-compulsive disorder: follow-up of 142 children and adolescents. *The British Journal of Psychiatry*, 197(2), 128-134.

National Institute for Health and Care Excellence. (2005) Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder. CG31. London: Author.

Pediatric OCD Treatment Study (POTS) Team. (2004). Cognitive-behavior therapy, sertraline, and their combination for children and adolescents with obsessive-compulsive disorder: The Pediatric OCD Treatment Study (POTS) randomized controlled trial. *JAMA: The Journal of the American Medical Association*, 292(16), 1969-1976.

Reaven, J. A., Blakeley-Smith, A., Culhane-Shelburne, K., & Hepburn, S. (2012). Group cognitive behavior therapy for children with high-functioning autism spectrum disorders and anxiety: a randomized trial. *Journal of Child Psychology and Psychiatry*, 53(4), 410-419.

Reaven, J. A., Blakeley-Smith, A., Nichols, S., Dasari, M., Flanigan, E., & Hepburn, S. (2009). Cognitive-Behavioral Group Treatment for Anxiety Symptoms in Children With High-Functioning Autism Spectrum Disorders A Pilot Study. *Focus on Autism and Other Developmental Disabilities*, 24(1), 27-37.

Reaven, J. A., & Hepburn, S. (2003). Cognitive-behavioral treatment of obsessive-compulsive disorder in a child with Asperger syndrome: a case report. *Autism*, 7(2), 145-164.

Russell, A. J., Jassi, A., Fullana, M. A., Mack, H., Johnston, K., Heyman, I., et al. (2013). Cognitive Behaviour Therapy for comorbid obsessive-compulsive disorder in high-functioning autism spectrum disorders: A randomised control trial. *Depression and Anxiety*, 30(8), 697-708.

Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., & Ort, S. I. (1997). Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(6), 844-852.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., et al. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, 40(11), 1228-1231.

Shafran, R., Frampton, I., Heyman, I., Reynolds, M., Teachman, B., & Rachman, S. (2003). The preliminary development of a new self-report measure for OCD in young people. *Journal of Adolescence*, 26(1), 137-142.

Shipman, D. L., Sheldrick, R. C., & Perrin, E. C. (2011). Quality of life in adolescents with autism spectrum disorders: reliability and validity of self-reports. *Journal of Developmental & Behavioral Pediatrics, 32*(2), 85.

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(8), 921-929.

Sofronoff, K., Attwood, T., & Hinton, S. (2005). A randomised controlled trial of a CBT intervention for anxiety in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry, 46* (11) 1152-1160.

Storch, E. A., Arnold, E. B., Lewin, A. B., Nadeau, J. M., Jones, A. M., De Nadai, A. S., et al. (2013). The effect of cognitive-behavioral therapy versus treatment as usual for anxiety in children with autism spectrum disorders: A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(2), 132-142.

Storch, E. A., Lewin, A. B., De Nadai, A. S., & Murphy, T. K. (2010). Defining treatment response and remission in obsessive-compulsive disorder: A signal detection analysis of the children's Yale-Brown obsessive compulsive scale. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(7), 708-717.

Storch, E. A., Merlo, L. J., Larson, M. J., Geffken, G. R., Lehmkuhl, H. D., Jacob, M. L., et al. (2008). Impact of Comorbidity on Cognitive-Behavioral Therapy Response in Pediatric Obsessive-Compulsive Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(5), 583-592.

Sze, K. M., & Wood, J. J. (2008). Enhancing CBT for the treatment of autism spectrum disorders and concurrent anxiety. *Behavioural and Cognitive Psychotherapy, 36*(04), 403-409.

Tani, P., Lindberg, N., Joukamaa, M., Nieminen-von Wendt, T., von Wendt, L., Appelberg, B., et al. (2004). Asperger syndrome, alexithymia and perception of sleep. *Neuropsychobiology*, 49(2), 64-70.

Uher, R., Heyman, I., Turner, C. M., & Shafran, R. (2008). Self-, parent-report and interview measures of obsessive-compulsive disorder in children and adolescents. *Journal of Anxiety Disorders*, 22(6), 979-990.

White, S. W., Oswald, D., Ollendick, T., & Scahill, L. (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review*, 29(3), 216-229.

Wood, J. J., Drahota, A., Sze, K., Har, K., Chiu, A., & Langer, D. A. (2009). Cognitive behavioral therapy for anxiety in children with autism spectrum disorders: a randomized, controlled trial. *Journal of Child Psychology and Psychiatry*, 50(3), 224-234.

Wood, J. J., & Gadow, K. D. (2010). Exploring the nature and function of anxiety in youth with autism spectrum disorders. *Clinical Psychology, Science and Practice*, 17(4), 281-292.

World Health Organization. (2008). ICD-10: International statistical classification of diseases and related health problems (10th Rev. Ed.). Geneva: Author.

Wu, M. S., McGuire, J. F., Arnold, E. B., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2013). Psychometric Properties of the Children's Yale-Brown Obsessive Compulsive Scale in Youth with Autism Spectrum Disorders and Obsessive-Compulsive Symptoms. *Child Psychiatry & Human Development*, 1-11.

Yeargin-Allsopp, M., Rice, C., Karapurkar, T., Doernberg, N., Boyle, C., & Murphy, C. (2003). Prevalence of autism in a US metropolitan area. *JAMA: The Journal of the American Medical Association*, 289(1), 49-55.